

# Menschzentrierte KI-Gestaltung

---

## Anschlussfähigkeit der Rollenkonzepte und Methodik von Constructive Technology Assess- ment und Value Sensitive Design

Zur Erlangung des akademischen Grades eines

DOKTORS DER PHILOSOPHIE (Dr.phil.)

von der KIT-Fakultät für Geistes- und Sozialwissenschaften  
des Karlsruher Instituts für Technologie (KIT)  
angenommene

DISSERTATION

von

Pascal Vetter

KIT-Dekan: Prof. Dr. Michael Mäs

1. Gutachter: Prof. Dr. Armin Grunwald
2. Gutachter: Prof. Dr. Karsten Wendland

Tag der mündlichen Prüfung: 10. März 2025





# Zusammenfassung

Systeme Künstlicher Intelligenz (KI) menschenzentriert zu gestalten erfordert Gestaltungsstrategien, die Verantwortungsträgern die Priorisierung menschlicher Fähigkeiten, Bedürfnisse und Werte in KI-Gestaltungsprozessen (KIGP) ermöglichen. Die Entwicklung entsprechender Gestaltungsmethoden ist eine Zielsetzung im Feld Human-centered AI (HCAI), dessen Realisierung eine zielführende Operationalisierung gestaltungstheoretischer Erkenntnisse erfordert, sowie eine bedachte Gewichtung von sozialen, ethischen, technischen und organisatorischen Gestaltungsfaktoren. Über die Navigation dieses multifaktoriellen Spannungsfelds herrscht im Feld Technikfolgenabschätzung (TA) ausgeprägtes Erfahrungswissen vor, dessen Aufgriff eine Erwägung für die Methodenentwicklung in HCAI lohnt. Die Anschlussfähigkeit von TA zu HCAI exploriert die Dissertation anhand der beiden Gestaltungsansätze Constructive Technology Assessment (CTA) und Value Sensitive Design (VSD). Die übergeordnete Forschungsfrage des Dissertationsprojekts lautet: Welche Erkenntnisse kann HCAI von den Ansätzen CTA und VSD aufgreifen, um KI-Systeme menschenzentriert zu gestalten? Zur Beantwortung dieser Frage fokussiert die Dissertation die verwendeten Rollenkonzepte in CTA und VSD, sowie die Methodik, die beide Ansätze zur Gestaltung technischer Systeme vorschlagen. Zunächst wird in Kapitel 2 die Gestaltungsphilosophie der menschenzentrierten Gestaltung (MZG) vorgestellt, die einen zentralen Ausgangspunkt konzeptueller und methodischer Überlegungen in HCAI bildet. Kapitel 3 gibt einen Überblick über das Feld HCAI und widmet sich den in HCAI diskutierten Weiterentwicklungsideen des MZG-Ansatzes, die sich bereits Mensch-Technik-Verhältnissen höherer Komplexität widmen. Kapitel 4 vollzieht dann einen Perspektivwechsel hin zur Auffassung von Technikgestaltung (TG) in TA und stellt zentrale Annahmen sowie die vorgeschlagenen Methoden der Ansätze CTA und VSD vor. Das Kapitel schließt mit den daraus resultierenden, zunächst mehrheitlich abstrakten Implikationen für die Ziele in HCAI. Kapitel 5 widmet sich anschließend den Rollenkonzepten, die in CTA und VSD verwendet werden, um menschliches Leben und Handeln in Gestaltungskontexten zu beschreiben und in Gestaltungsmethoden zu operationalisieren. Ziel der Ergründung der Bedeutungen, Ursprünge und blinden Flecke der Rollenkonzepte ist nachzuvollziehen, welche Interpretationen von Menschzentriertheit in CTA und VSD zum Tragen kommen, und wie sich diese auf die vorgeschlagenen Methoden übertragen. Daraus resultierende Schlussfolgerungen für HCAI werden am Ende des Kapitels diskutiert. Kapitel 6 wendet sich der Gestaltungspraxis zu und beschreibt die Ergebnisse einer Interviewstudie, in der KI-Experten aus der Gestaltungspraxis die Praxistauglichkeit der CTA- und VSD-Methoden beurteilen. Zentrale Leitfragen der Expertenevaluation sind, inwieweit die Methoden den Anforderungen von KIGP in der Praxis gerecht werden können und inwieweit sie Verantwortungsträgern das Treffen belastbarer Gestaltungsentscheidungen ermöglichen, da dies zentrale Kriterien für HCAI-Methoden darstellen. Die differenzierten Einschätzungen der KI-Experten werden mithilfe einer SWOT-Analyse dargestellt, deren Ergebnisse die Diskussion weiterer Schlussfolgerungen für HCAI ermöglichen, sowie die Formulierung von Weiterentwicklungsempfehlungen für die Methoden von CTA und VSD. In Kapitel 7 wird zusammenfassend darüber reflektiert, was CTA und VSD für HCAI (nicht) bereitstellen können und unter welchen Bedingungen eine interdisziplinäre Zusammenarbeit der Communities HCAI und TA zielführend sein kann.



# Abstract

Designing and developing artificial intelligence (AI) systems in a human-centered way requires design strategies that enable decision-makers to prioritize human capabilities, needs and values in AI design processes. The development of corresponding design methods is an objective in the field of Human-centered AI (HCAI), the realization of which requires a goal-oriented operationalization of design theory findings as well as a considered weighting of social, ethical, technical and organizational design factors. The field of Technology Assessment (TA) has a wealth of experience in navigating this multifactorial field of possible tensions and conflicts, which is worth considering for the development of methods in HCAI. The dissertation explores the compatibility of TA with HCAI through closely examining the two design approaches Constructive Technology Assessment (CTA) and Value Sensitive Design (VSD). The central research question of the dissertation project is: What insights can HCAI draw from the CTA and VSD approaches in order to design AI systems in a human-centered way? To answer this question, the dissertation focuses on the role concepts used in CTA and VSD, as well as the methods that both approaches propose for designing and developing technical systems. Chapter 2 begins with an introduction to the design philosophy of Human-centered Design (HCD), which forms a central starting point for conceptual and methodological considerations in HCAI. Chapter 3 provides an overview of the field of HCAI and presents the ideas for further development of the HCD approach discussed in HCAI, which already address human-technology-relationships of higher degrees of complexity. Chapter 4 then shifts the perspective towards the understanding of technology design in TA and presents central assumptions as well as the proposed design methods of CTA and VSD. The chapter concludes with the resulting, initially mostly abstract implications for the objectives in HCAI. Chapter 5 is then devoted to the role concepts used in CTA and VSD to describe forms of human life and action in design contexts and to operationalize them in design methods. The aim of exploring the meanings, origins and blind spots of these role concepts is to understand which interpretations of human-centeredness come into play in CTA and VSD, and how these translate into the respective design methods of the both approaches. The chapter ends by discussing the resulting conclusions for HCAI. Chapter 6 then turns towards design practice and describes the results of an interview study in which AI experts assess the practicability of the CTA and VSD methods. The central questions of the expert evaluation focus on the extent to which the methods can meet the requirements of AI design processes in practice and to the extent to which they enable decision-makers to make reliable design decisions, as these present central criteria for HCAI methods. The differentiated assessments of the AI experts are presented with the help of a SWOT analysis, the results of which enable the discussion of further conclusions for HCAI and the formulation of recommendations for the further development of the CTA- and VSD-methods. Chapter 7 summarizes what CTA and VSD can (not) provide for HCAI and under what conditions interdisciplinary cooperation between the HCAI and TA communities can be expedient.

# Vorbemerkungen

Die Dissertation widmet sich einem interdisziplinären Thema, weshalb im Sprachgebrauch der Arbeit disziplinäre Konventionen aufgegriffen und wiedergegeben werden. Die Verwendung solcher Konventionen ist erforderlich, um Positionen, Gedankengänge und Argumentationspfade nachzuzeichnen, sowie um die Eigenschaften und Besonderheiten verschiedener Wissenstraditionen und Wissenskulturen abzubilden. Eine solche Konvention besteht darin, zum Zwecke der Darlegung verschiedener Perspektiven auf und Interpretationen von Beziehungen zwischen Menschen und Technik, Unterscheidungen zwischen menschlichem Leben und Handeln und den Funktionen und Abläufen technischer Systeme vorzunehmen. Dazu findet häufig der generische Plural „Menschen“ sowie der generische Singular „Mensch“ Verwendung. Als Autor liegt mir nichts ferner, als Wertungen von Lebensentwürfen und Lebensstilen vorzunehmen, weshalb bei der Verwendung dieser Bezeichnungen ausdrücklich sämtliche Kulturen, Religionen und Ethnien angesprochen sind. Zur Förderung der Lesbarkeit und Nachvollziehbarkeit verzichte ich zudem darauf, bei Personenbezeichnungen sowohl die weibliche als auch die männliche und diverse Form zu nennen. Das generische Maskulinum adressiert ausdrücklich alle Leserinnen und Leser und gilt, sofern nicht anders gekennzeichnet, für alle Geschlechter.

Eine weitere sprachliche Konvention die ich in der Dissertation aufgreife, ist die Segmentierung von Technikgestaltungsprozessen in die Gestaltung und die Nutzung von Technik. In den Wissenskulturen, die sich mit Technik befassen, erfährt der Begriff der Gestaltung eine sehr hohe Bedeutungspluralität, insbesondere durch seine inhaltlichen Überschneidungen mit den Begriffen der Konstruktion, der Entwicklung und dem Design von Technik (Hellige, 1995). Die Bezeichnung der Entwicklung von Technik findet dabei besonders in den Communities Verwendung, die sich der Erfindung und Realisierung digitaler Systeme widmen. Die Tätigkeit der Entwicklung von Software-, Hardware- und KI-Systemen erfordert ausgeprägtes technisches Knowhow, jedoch sind dafür auch Fähigkeiten wie Einfallsreichtum, Ausdauer und Flexibilität notwendig. Um dieser Vielfalt menschlicher Fähigkeiten auf angemessene Weise Rechnung zu tragen, ziehe ich es vor, in der Dissertation vornehmlich von der Gestaltung technischer Systeme zu sprechen. Zur Konvention, Menschen in Kontexten von Technikgestaltung als „Nutzer“ zu konzipieren, muss ich darüber hinaus auf den Diskurs hinweisen, der die dehumanisierenden Tendenzen dieser Tradition kritisiert. Die zentralen Kritikpunkte fasst Donald Norman treffend zusammen:

*„Psychologists depersonalize the people they study by calling them “subjects.” We depersonalize the people we study by calling them “users.” Both terms are derogatory. They take us away from our primary mission: to help people. Power to the people, I say, to repurpose an old phrase. People. Human beings. That’s what our discipline is really about”* (Norman, 2006, S. 49).

Im Verlauf der Dissertation werde ich Inhalte dieses Diskurses aufgreifen und referenzieren, da sie für die Fragestellung der Arbeit besondere Relevanz erfahren. Wenn ich selbst die Bezeichnung Nutzer verwende, tue ich dies vorwiegend, um etablierten sprachlichen Konventionen der Communities Rechnung zu tragen, die für die inhaltlichen Auseinandersetzungen der Dissertation relevant sind.

# Danksagung

Das Verfassen dieser Dissertation war für mich auf viele verschiedene Weisen eine prägende und bereichernde Erfahrung, die jedoch ohne die Unterstützung der Menschen in meinem Umfeld kaum möglich gewesen wäre. Ich möchte deshalb an dieser Stelle die Möglichkeit ergreifen den Menschen zu danken, die mich in meiner Promotionszeit in besonderer Weise begleitet und unterstützt haben. Mein herzlicher Dank gebührt in erster Linie meinen beiden Betreuern Prof. Armin Grunwald und Prof. Karsten Wendland, die zu jeder Zeit ein offenes Ohr für meine Fragen und Anliegen hatten und durch ihre unzähligen wertvollen Anregungen und Hinweise einen wesentlichen Beitrag zum Gelingen dieser Arbeit geleistet haben. Die Kombination aus Freiheit und Anleitung die sie mir gewährten, war ein Schlüsselement für die erfolgreiche Bearbeitung des Dissertationsthemas.

Ich hatte darüber hinaus das große Glück, während der Promotionszeit Menschen zu begegnen, die erst zu Kollegen und schließlich zu Freunden wurden. Durch ihr kontinuierliches Feedback erhielt ich wertvolle Impulse und konnte Gedanken entwickeln, die von entscheidender Bedeutung für die Bearbeitung eines interdisziplinären Promotionsthemas sind. Herzlich danken möchte ich in diesem Zusammenhang in besonderer Weise Anna Keitemeier, die mir in zahlreichen inspirierenden (Streit-)Gesprächen dabei geholfen hat, die Schönheit und Vielfalt kulturwissenschaftlicher Arbeit zu erfahren. Auch danke ich Hüseyin Ersoy herzlich, der mir oftmals selbstlos seine Zeit und seine Aufmerksamkeit geschenkt hat und es mir durch seine weisen Ratschläge ermöglichte, auch die Täler der Promotionszeit erfolgreich zu navigieren. Darüber hinaus gilt mein herzlicher Dank auch denjenigen Freunden, die mich von Beginn an unterstützt und bestärkt haben. Hier danke ich herzlich Rebecca Gerstenberg und Gregor Lenhard, die beide besonders zu Beginn in regelmäßigen Gesprächen dazu beigetragen haben, dass ich meinen Fokus bewahren kann. Auch Andreas Mayer danke ich in diesem Zusammenhang herzlich, da er mir in unzähligen Gesprächen dabei half, die Herausforderungen, die das Verfassen einer Dissertation mit sich bringen kann, aus verschiedenen Perspektiven zu betrachten, um auch in kritischen Phasen die Ruhe bewahren zu können.

Auf familiärer Seite möchte ich zudem besonders meinem Bruder Marcel Vetter herzlich danken, der mich immer wieder geerdet hat und mit dessen Unterstützung es mir möglich wurde, auch in arbeitsintensiven Phasen erholsame und freudvolle Stunden zu verbringen.

Karlsruhe, im Juni 2024

*Pascal Vetter*

# Abkürzungsverzeichnis

ANT	Actor-Network-Theory
CSCW	Computer Supported Cooperative Work
CTA	Constructive Technology Assessment
HCAI	Human-centered AI
HCD	Human-centered Design
HCI	Human-Computer Interaction
IT	Informationstechnologie
KI	Künstliche Intelligenz
KIGP	KI-Gestaltungsprozesse
MSEP	Modellierung von Softwareentwicklungsprozessen
MZG	Menschzentrierte Gestaltung
MZKG	Menschzentrierte KI-Gestaltung
PD	Participatory Design
SCOT	Social Construction of Technology
SE	Software Engineering
STS	Science and Technology Studies
TA	Technikfolgenabschätzung
TG	Technikgestaltung
UCD	User-centered Design
UI	User Interface
UX	User Experience
VSD	Value Sensitive Design

# Inhalt

1	Einleitung .....	1
1.1	Konzeptionelle und methodische Herausforderungen .....	2
1.2	Zur Beitragsfähigkeit der Technikfolgenabschätzung .....	4
1.3	Struktur der Arbeit .....	6
2	Menschzentrierte Gestaltung .....	10
2.1	Einführung .....	10
2.2	Historie und Grundgedanken .....	13
2.2.1	Kritik unreflektierter Computereuphorie (Joseph Weizenbaum) .....	13
2.2.2	Integratives Verständnis von Gestaltung und Nutzung (Christiane Floyd).....	15
2.2.3	Anbindung von Erkenntnissen der Psychologie und der Designforschung (Donald Norman) ..	19
2.2.4	Fokussierung von Nutzerpraktiken und Nutzungskontexten (Lucy Suchman) .....	22
2.3	Gestaltungsprinzipien .....	24
2.3.1	Erfassung des Nutzungskontexts .....	25
2.3.2	Nutzerbeteiligung .....	26
2.3.3	Iterative Gestaltungsprozesse .....	27
2.3.4	Interdisziplinäre Gestaltungsteams .....	27
2.4	Zusammenfassung .....	28
3	Menschzentrierte KI-Gestaltung .....	29
3.1	Einführung .....	29
3.2	Überblick der Forschungslandschaft .....	31
3.2.1	Einrichtungen .....	31
3.2.2	Themenfelder .....	33
3.3	Erweiterte Modelle menschzentrierter Gestaltung .....	35
3.3.1	Einordnung in KI-Lebenszyklen.....	36
3.3.2	Information des politischen Geschehens.....	37
3.3.3	Implementierung moderner Werkzeuge .....	38
3.3.4	Balancierung von Autonomie und Automatisierung.....	40
3.4	Trends .....	42
3.4.1	Realisierung ko-kreativer und verantwortungsvoller KI.....	42
3.4.2	Reflektierter Umgang mit Konzeptionen zu menschlichem Leben und Handeln.....	43
3.4.3	Zur Bedeutung der Humanities für HCAI .....	44
3.4.4	Verantwortungsbereiche von HCAI .....	45
3.5	Zusammenfassung .....	46
4	Gestaltungsvermögen in Technikfolgenabschätzung .....	47
4.1	Einführung .....	47
4.2	Constructive Technology Assessment .....	49
4.2.1	Zentrale Annahmen und Konzepte.....	50
4.2.2	Vorgeschlagene Methodik .....	51
4.3	Value Sensitive Design .....	56
4.3.1	Zentrale Annahmen und Konzepte.....	58
4.3.2	Vorgeschlagene Methodik .....	61
4.4	Zwischenbetrachtung .....	69
4.4.1	Abstraktionsebene und Blick auf Gestaltungsprozesse .....	69
4.4.2	Menschorientierung.....	71
4.4.3	Methodische Vagheit .....	73
4.4.4	Erfolgskritische Aspekte der CTA-Methodik.....	74
4.4.5	Operationalisierbarkeit für die KI-Gestaltung .....	75
4.5	Zusammenfassung .....	76

5	Rollenkonzepte in CTA und VSD .....	77
5.1	Einführung .....	77
5.2	Rollenverständnisse im Gestaltungskontext.....	79
5.2.1	Rollenverständnis von Ralph Linton.....	81
5.2.2	Rollenverständnis in der Modellierung von Softwareentwicklungsprozessen.....	83
5.2.3	Integratives Verständnis von Rollenkonzept .....	85
5.3	Vorgehensweise .....	87
5.3.1	Durchführung .....	87
5.3.2	Auswertung.....	90
5.4	Verwendete Rollenkonzepte .....	91
5.4.1	Die Involvierten .....	92
5.4.2	Die informierten Einflussnehmer.....	99
5.4.3	Die Unbeteiligten und Betroffenen .....	108
5.4.4	Die Gestalter .....	112
5.4.5	Die intervenierenden Prozessbegleiter.....	118
5.4.6	Die Nutzer .....	121
5.4.7	Die Technikenthusiasten.....	123
5.4.8	Die Technikgegner .....	125
5.5	Ursprünge der Rollenkonzepte.....	126
5.5.1	Die Involvierten .....	127
5.5.2	Die informierten Einflussnehmer.....	131
5.5.3	Die Unbeteiligten und Betroffenen .....	134
5.5.4	Die Gestalter .....	137
5.5.5	Die intervenierenden Prozessbegleiter.....	138
5.5.6	Die Nutzer .....	139
5.5.7	Die Technikenthusiasten und die Technikgegner.....	140
5.6	Blinde Flecke der Rollenkonzepte.....	143
5.6.1	Abstraktion vs. Spezifität .....	144
5.6.2	Die Rolle des Individuums.....	145
5.6.3	Risiken des Nutzerrollenkonzepts.....	147
5.6.4	Fehlende Betrachtung von Managementrollenkonzepten.....	149
5.6.5	Zum Auswirkungsanspruch von CTA agents .....	150
5.6.6	Differenzierung der Motivationen von Enactors .....	153
5.6.7	Unvollständige Auseinandersetzung mit Technikgegnern.....	154
5.7	Diskussion .....	156
5.7.1	Erweiterung von Denkhaltungen und Perspektiven .....	157
5.7.2	Interpretationen identisch bezeichneter Rollenkonzepte.....	160
5.7.3	Die Bedeutung der Unbeteiligten und Betroffenen .....	163
5.7.4	Rollenkonzepte des Data Managements.....	165
5.7.5	Strukturelle Unterschiede zwischen der Gestaltung von Software- und KI-Systemen.....	166
5.7.6	Die Verwendung von Rollenkonzeptpaaren .....	167
5.7.7	Mehrwert für die Untersuchung ethischer Aspekte .....	168
5.7.8	Verwendung verschiedener Abstraktionsebenen.....	170
5.8	Zusammenfassung.....	172
6	Methodik von CTA und VSD .....	174
6.1	Einführung .....	174
6.2	Vorgehensweise .....	176
6.2.1	Vorbereitung .....	176
6.2.2	Durchführung .....	180
6.2.3	Auswertung.....	182
6.3	Stärken.....	183
6.3.1	Prozess- und Ergebnisorientierung.....	183
6.3.2	Befähigung zu Reflexion und Innovation.....	186
6.3.3	Partizipation von Stakeholdern .....	190

6.4	Schwächen.....	192
6.4.1	Prozess- und Ergebnisorientierung.....	192
6.4.2	Befähigung zur Reflexion .....	195
6.4.3	Partizipation von Stakeholdern .....	197
6.5	Chancen.....	199
6.5.1	Prozess- und Ergebnisorientierung.....	199
6.5.2	Wissenstransfer und Werteorientierung .....	201
6.5.3	Kompatibilität und Variabilität.....	202
6.5.4	Befähigung zur Reflexion und Folgenabschätzung.....	203
6.6	Risiken.....	207
6.6.1	Prozess- und Ergebnisorientierung.....	207
6.6.2	Reflektierter Umgang mit Verantwortung und Befangenheit .....	211
6.6.3	Partizipation von Stakeholdern .....	212
6.7	Diskussion .....	213
6.7.1	Zur Funktion und den Inhalten der Storyboards .....	213
6.7.2	Transfer von Reflexionserkenntnissen .....	214
6.7.3	Steigerung von Ausbildungsaktivitäten.....	215
6.7.4	Aufgreifen bewährter Mixed-Method-Ansätze .....	216
6.7.5	Gestaltung und Projektmanagement .....	217
6.7.6	Klärung von Verantwortungsbereichen .....	217
6.7.7	Flexibilität und Agilität .....	219
6.7.8	Verankerung der Folgenabschätzung in der Praxis .....	219
6.8	Zusammenfassung .....	220
7	Folgerungen .....	221
8	Literaturverzeichnis .....	224
9	Abbildungsverzeichnis.....	253
10	Tabellenverzeichnis.....	256
11	Anhang .....	257





# 1 Einleitung

Die Zielsetzungen und Vorschläge der menschenzentrierten Gestaltung (MZG) digitaler Systeme erfahren im Jahr 2024 viel Zuspruch (Progress Software Corporation, 2024). Für viele Gestalter, die konzipierend, entwickelnd oder organisierend tätig sind, ist der Leitgedanke informationstechnische (IT) Systeme für menschliche Bedürfnisse und Anforderungen zu gestalten, zu einem festen Bestandteil ihres Berufsethos geworden. Darüber hinaus haben sich die Gedanken und Konzepte der MZG auch auf die Organisationsentwicklung übertragen. Es hat sich gezeigt, dass die in der MZG verankerten Prozesskonzepte auch dazu geeignet sind, übergreifende Strukturen von Organisationen sozialförderlich auszurichten und damit einen signifikanten Wettbewerbsvorteil zu nutzen (Rae, 2016; Sheppard et al., 2018). Durch verständliches Systemdesign, das Effektivität, Effizienz und Zufriedenstellung anstrebt (Usability; Lewis, 2014; Nielsen, 1994), lernen Organisationen sich das Vertrauen von Kunden, Mitarbeitern und Geschäftspartnern langfristig zu sichern und so die Risiken finanzieller Fehlinvestitionen abzusenken. Orientierung zur Umsetzung menschenzentrierter Organisationsformen bietet ihnen beispielsweise die Norm DIN EN ISO 27500 (DIN EN ISO 27500:2017-07, 2017), die dazu beiträgt, dass menschenorientierte Organisationsstrukturen auf internationaler Ebene an Relevanz gewinnen (Townsend & Romme, 2024). Dies sind begrüßenswerte Entwicklungen, da sie zeigen, dass die jahrzehntelangen Bemühungen von vielen Wissenschaftlern verschiedener Disziplinen ihren Weg in die Gestaltungs- und Organisationspraxis gefunden haben, und über digitale Produkte und neue Organisationsformen auch den Weg in die Mitte des gesellschaftlichen Lebens finden. Angesichts anhaltender sozialer und technischer Entwicklungen werfen die Konzepte und Ideen der MZG jedoch neue Fragen auf. Auf der technisch-organisationalen Ebene lautet eine zentrale Fragestellung: Wie können die Zielsetzungen der MZG auf die Gestaltung von Systemen angewandt werden, die moderne Algorithmen und Verfahren digitaler Technologien verwenden? Damit eng verknüpft ist die Frage, welche Vorstellungen von menschlichem Leben und Handeln in der Gestaltung aufgegriffen und verwendet werden können (und müssen), um menschliches Leben und Handeln auf angemessene Weise zu repräsentieren. Diese beiden Fragen stehen im Mittelpunkt des gesellschaftlichen und wissenschaftlichen Diskurses zur Gestaltung und zum Einsatz von Systemen Künstlicher Intelligenz (KI). Die Motivation der Dissertation besteht darin, einen Beitrag zur Beantwortung dieser Fragen zu leisten. Als Autor interessiert mich dabei besonders, wie angesichts der technischen Möglichkeiten die KI eröffnet, die Gedanken und Ideen einer menschenzentrierten Gestaltungsphilosophie gefasst werden können, und wie es gelingen kann, diese Erkenntnisse möglichst verlustfrei und zielführend in die Gestaltungspraxis zu überführen.

In den Debatten zur KI-Gestaltung herrscht weitgehend Einigkeit darüber, dass KI-Systeme menschenzentriert gestaltet werden sollen. Viele Debatten verharren jedoch bei dieser Forderung und unterlassen das Nachgehen der daraus resultierenden Frage, wie eine menschenzentrierte Gestaltung von KI-Systemen (MZKG) konkret realisiert werden kann. Bei der Auseinandersetzung mit dieser Frage wird rasch deutlich, dass das Erarbeiten von Antwort-

ten das Einnehmen einer ganzheitlichen Perspektive auf KI-Gestaltung erfordert, die eine integrative Berücksichtigung von technischen, organisatorischen, sozialen und ethischen Gestaltungsfaktoren ermöglicht. Zudem müssen Gestaltungstheorie und Gestaltungspraxis eng zusammenarbeiten – es gilt Werkzeuge und Hilfestellungen zu entwickeln, die es Gestaltungspraktikern erlauben, die von der Gesellschaft an sie gerichtete Forderung menschenzentrierter KI-Systeme zu realisieren und die gleichzeitig auf die Anforderungen eingehen, mit denen sie sich in der Gestaltungspraxis bei der Umsetzung dieser Aufgabe konfrontiert sehen. Über den Zusammenschluss gestaltungstheoretischer und gestaltungspraktischer Erkenntnisse verfügen verschiedene Disziplinen über ausgewiesene Expertise, die zur Bewältigung dieser Aufgabe aufgegriffen werden können. Zwei solcher Disziplinen stellen die Human-Computer Interaction (HCI) und die Technikfolgenabschätzung (TA) dar. Aus der HCI, der Heimatdisziplin des MZG-Ansatzes, hat sich in den vergangenen Jahren die Community Human-centered AI (HCAI) herausgebildet, die sich intensiv mit den KI-induzierten Gestaltungsfragen auseinandersetzt. Das Feld TA widmet sich traditionell der Abschätzung von gesellschaftlichen Chancen und Risiken technologischer Entwicklungen, und vertritt in den Auseinandersetzungen mit KI insbesondere geistes- und sozialwissenschaftliche Positionen, die weitreichende Zeithorizonte und komplexere soziotechnische Ursache-Wirkungs-Ketten antizipieren. In beiden Feldern ist dadurch die benötigte interdisziplinäre Expertise vorhanden, deren Zusammenführung eine realisierbare MZKG erfordert. Vor dem Hintergrund konkreter Fragestellungen zur MZKG explore ich in der Dissertation die Anschlussfähigkeit von TA zu HCAI anhand der beiden Gestaltungsansätze Constructive Technology Assessment (CTA) und Value Sensitive Design (VSD). Für meine Darlegungen greife ich auf meine Expertise in HCI und meine aufgebaute Expertise in TA zurück.

### 1.1 Konzeptionelle und methodische Herausforderungen

Ein Strang des HCAI-Diskurses beschäftigt sich mit den übergeordneten Annahmen zu menschlichem Leben und Handeln, die in KI-Gestaltungsprozessen (KIGP) zum Tragen kommen und Verwendung finden (Auernhammer, 2020; Bingley et al., 2023; Capel & Brereton, 2023; Inkpen et al., 2019; Subramonyam et al., 2022; Xu, Dainoff, et al., 2023). In vielen Bereichen der KI-Gestaltung wird dabei hinsichtlich des Verhältnisses von Mensch und KI-System auf abstrakter Ebene eine grundsätzliche Rollenaufteilung vorgenommen, die zwischen Gestaltern und Nutzern von KI-Systemen unterscheidet. Diese zweiseitige Rollenaufteilung hat sich in der Vergangenheit in Gestaltungskontexten von IT-Systemen bewährt, um der inhaltlichen und organisatorischen Komplexität von Gestaltungsprozessen zu begegnen, und wird entsprechend auf KIGP übertragen. Damit einher geht der Rückgriff auf traditionell etablierte Rollenkonzepte, der pragmatische Vorzüge für die Vorgänge in KIGP besitzt, der jedoch auch die damit verbundenen Risiken einer mehr einseitigen Definition von Rollen und Rollenattributen durch Gestalter, Projektmanager oder anderweitige Verantwortungsträger übernimmt. Ergebnis kann die Verwendung stellenweise wenig reflektierter Rollenkonzepte wie das eines Dumbest Assumable Users

(Spiekermann, 2010) sein, die soziales Verhalten unterkomplex und fehlinterpretierbar abbilden und deren Verwendung dazu führen kann, dass die Orientierung an (tatsächlichen) menschlichen Verhaltensweisen und Bedürfnissen – trotz explizit formuliertem Anspruch – in Gestaltungsprozessen verloren geht (Wendland, 2021b). Für die Zielsetzungen der MZKG ist die sorgfältige Analyse und bedachte Verwendung von Rollenkonzepten von zentraler Bedeutung, da Rollenkonzepte durch ihre Funktion als gedankliche und methodische Referenzkonstrukte in Gestaltungsansätzen entscheidend auf die Menschorientierung in Gestaltungsprozessen und schließlich den daraus resultierenden KI-Systemen starken Einfluss nehmen. Angesichts der weitflächigen Einsatzmöglichkeiten von KI-Systemen gewinnt dieser Aspekt zusätzlich an Relevanz. Die Analyse verwendeter Rollenkonzepte wird nach meiner Einschätzung in der MZKG-Debatte bislang jedoch stark vernachlässigt. Diese Forschungslücke adressiere ich in der Dissertation durch eine Analyse der Rollenkonzepte, die in den beiden im Fokus stehenden Gestaltungsansätzen CTA und VSD Verwendung finden.

Mit den Gestaltungsansätzen ist ein zweiter Strang des HCAI-Diskurses angesprochen, der gegenwärtig besonders für die Gestaltungspraxis an Bedeutung gewinnt. Diese weist einen Bedarf an Gestaltungsframeworks, besonders an konkreten Gestaltungsmethoden auf, die eine menschenzentrierte Gestaltung von KI-Systemen unter den in der Gestaltungspraxis vorherrschenden Bedingungen ermöglichen (Deng et al., 2023; Feng et al., 2023; Widder & Nafus, 2023; Windl et al., 2022; Yang et al., 2020; Yildirim et al., 2023). Der Wissenstransfer in die Gestaltungspraxis vollzieht sich jedoch nur langsam. Erste Operationalisierungsleistungen sind in Form von Gestaltungsprinzipien (z.B. Cronholm & Göbel, 2022; Dimitriadis et al., 2021; Interaction Design Foundation, 2024; Pacaux-Lemoine et al., 2017) und Gestaltungsrichtlinien (z.B. Amershi et al., 2019; H. Chen et al., 2022; Shneiderman, 2020) unterschiedlicher Ausprägung erbracht worden, erweisen sich in der Gestaltungspraxis jedoch oftmals als zu optimistisch oder zu vage, wodurch sie von Gestaltungspraktikern nur bedingt als belastbare Hilfsmittel betrachtet werden. Ferner sind verschiedene MZKG-Frameworks und Methoden entwickelt worden, die den Weg in die Gestaltungspraxis bislang nicht gefunden haben, und über die entsprechend wenig Erkenntnisse hinsichtlich ihrer Praxistauglichkeit vorliegen. Aufgrund dieser Sachlage stimme ich den Einschätzungen zu, die einen Bedarf zur Steigerung des bilateralen Wissensaustauschs zwischen Gestaltungstheorie und Gestaltungspraxis identifiziert haben (Hartikainen et al. 2022; Mittelstadt 2019; Ehsan et al. 2021; Braun et al. 2021; Dignum und Dignum 2020), besonders im Hinblick auf die Rückführung von Praxiswissen in die (Weiter-)Entwicklung von Gestaltungsansätzen. Diese Forschungslücke adressiere ich in der Dissertation durch eine Expertenevaluation der Methoden von CTA und VSD.

## 1.2 Zur Beitragsfähigkeit der Technikfolgenabschätzung

Eine Strategie für den Wissenserwerb zu interdisziplinären Fragestellungen besteht darin, in verschiedene Felder zu blicken, die sich mit relevanten Zusammenhängen und Teilaspekten einer Frage auseinandersetzen oder die über relevante Expertisen verfügen, die wertvolle Einsichten beitragen können. Für die HCAI-Fragestellungen zu reflektierten Rollenkonzepten und praxistauglichen Methoden für die MZKG stellt die TA ein Feld dar, das im Ansatz beide Kriterien erfüllt und deshalb eine genauere Betrachtung erstrebenswert macht. In TA-Arbeiten wird die KI-Technologie vorwiegend unter verschiedenen Gesichtspunkten ihrer gesellschaftlichen Auswirkungen untersucht, weshalb die Inhalte der Arbeiten mit KI-Bezügen weit gestreut sind. Risiken von KI-Anwendungen zur automatisierten Verarbeitung von Sprache (Albrecht, 2023; Schreiber & Ohly, 2024) oder zur Generierung realistisch wirkender Medieninhalte (Deepfakes; European Parliament et al., 2021) können dabei gleichermaßen im Mittelpunkt stehen wie die Diskriminierungsrisiken von in KI-Systemen verwendeten Algorithmen (Orwat, 2019) oder die Entmystifizierung der mit KI verbundenen Erwartungen und Befürchtungen (Humm et al., 2021; Wendland, 2021a). Auch wenn KI nur eine von vielen Technologien ist, die in TA einen Untersuchungsgegenstand darstellt, erkenne ich in den Motivationen und Zielsetzungen der Auseinandersetzungen mit KI Parallelen zu jenen in HCAI. Die für die Fragestellung der Dissertation wesentlichen Parallelen skizziere ich nachfolgend kurz, und gehe dabei auch auf die relevanten Unterschiede der Wissenstraditionen und Wissenskulturen beider Felder ein.

Grundsätzlich besitzen beide Felder eine interdisziplinär geprägte Geschichte, in der die Betrachtung von Technik und ihrer Gestaltung durch einen Verbund verschiedener disziplinärer Perspektiven erfolgt. In HCAI werden aufgrund seiner Entstehung aus HCI mehrheitlich Erkenntnisse der Informatik, Psychologie, Soziologie und des Designs verknüpft, um KI-Systeme zu gestalten. Im Vergleich dazu liegt in TA der Perspektivenschwerpunkt zu KI auf Aspekten der Technikethik (Berberich, 2019; Grunwald, 2020; Heil, 2021; Weydner-Volkmann, 2021), der Techniksoziologie und der Technikphilosophie. In TA stehen beispielsweise Fragen im Vordergrund, wie KI-Systeme moralisch bewertet werden können (Seng, 2019) und welche Auswirkungen die erlebbaren technischen Fortschritte von KI auf das Selbstverständnis des Menschen haben (Grunwald, 2021; Quinn, 2021). Zwar werden in derartigen TA-Arbeiten Fragen zur KI-Gestaltung mehr auf einer Metaebene behandelt, jedoch werden spätestens beim Blick auf Fragestellungen zur Gestaltung von Schnittstellen für die Interaktion zwischen Mensch und KI-System die Bezüge zu Fragestellungen in HCAI erkennbar (z.B. Liggieri & Müller, 2019). In beiden Feldern liegt der Fokus der Betrachtung des Verhältnisses von Mensch und Technik auf der Seite der Menschen, deren Schutz und Wohlbefinden Aktivitäten der Technikgestaltung (TG) oberste Priorität einräumen sollen. KI-Systeme sollen Menschen helfen die von ihnen gewünschten Aufgaben durchzuführen, sie vor Schäden jeglicher Art bewahren und somit im positiv konnotierten Sinne zu einem „human flourishing“ (Stahl et al., 2021) auf gesellschaftlicher Ebene beitragen. Ein weiterer Anknüpfungspunkt zwischen beiden Feldern liegt in der problem-

orientierten Auseinandersetzung mit Fragestellungen zu KI, wodurch häufig die Operationalisierung theoretischer Erkenntnisse zur Bewältigung realweltlicher Herausforderungen im Vordergrund steht. Aufgrund dieses Umstandes verfügen beide Felder über viel Erfahrungswissen zum Umgang mit den oftmals konfligierenden Interessen von Gestaltern und Nutzern von Technik. Mit Blick auf Folgen von KI besteht in TA die Aufgabe oftmals darin, die Auswirkungen der Gestaltung und des Einsatzes von KI-Systemen zu antizipieren, auf übergreifende soziale Zusammenhänge abzubilden und dadurch möglicherweise auftretende Schäden für den Menschen abzuwenden. In HCAI steht hingegen primär das Heben der als erstrebenswert betrachteten Potenziale der KI-Technologie im Vordergrund, die in konkreten, als relevant identifizierten sozialen Kontexten von Bedeutung sind. In TA wird deshalb häufig auf der Abstraktionsebene der gesellschaftlichen Relevanz argumentiert, wohingegen in HCI traditionell kaum über einen bestimmten Nutzungskontext hinaus argumentiert wird und sich diese mehr ganzheitliche Diskurstradition erst allmählich in HCAI zu entfalten beginnt. Damit zusammenhängend steht in TA die Antizipation und Prävention möglicher nicht-intendierter Folgen von KI-Systemen mehr im Vordergrund, während in HCAI das Hauptinteresse der Förderung und Realisierung intendierter Folgen gilt. Dies verdeutlicht die grundlegend unterschiedliche Auffassung von TG an sich sowie das grundlegend unterschiedliche Selbstverständnis hinsichtlich der Aktivität in Technikgestaltungsprozessen: Im Feld TA ist ein Selbstverständnis in Bezug auf TG vorhanden, das sich dadurch auszeichnet, Technik mehr indirekt und überwiegend durch Regulierungsmaßnahmen zu „gestalten“, während in HCAI das konkrete und aktive Gestaltungsvorhaben der KI-Technologie auf technischer Ebene stark im Vordergrund steht. In beiden Feldern erfahren zudem auf methodischer Ebene angewandte partizipative Formate einen hohen Stellenwert, was sich sowohl in der Methodenlandschaft widerspiegelt, wie im Wissen um die Bedeutung der Auseinandersetzung mit scheinbar geringfügig relevanten Stakeholdergruppen in Gestaltungsprozessen.

Diese Ausgangslage motiviert mich zur Durchführung einer explorativen Untersuchung der Anschlussfähigkeit von TA-Aspekten zu HCAI anhand konkreter Gesichtspunkte. TA-Aspekte haben bislang nur wenig Eingang in HCAI-Betrachtungen gefunden, die dort aufgrund der angestiegenen inhaltlichen Komplexität um gesellschaftliche, politische und philosophische Fragestellungen zielführende Beiträge zu einer mehr ganzheitlichen Betrachtung von Gestaltungsprozessen leisten können. Offen ist die Frage, inwieweit TA-Erkenntnisse mit ihren traditionell geringen Ambitionen zur aktiven Beteiligung an Gestaltungsaktivitäten und der Übernahme von Verantwortung für Gestaltungsentscheidungen konkrete Beiträge zu HCAI-Fragestellungen leisten können. Ein willkommener Beitrag für die derzeitige HCAI-Debatte sind nach meiner Einschätzung Hinweise auf Ansätze und Erfahrungswerte, die einen Übertrag von Erkenntnissen zur erstrebenswerten Gestaltung von Mensch-KI-Verhältnissen in die operative Gestaltungspraxis ermöglichen. Vor dem Hintergrund der Fragestellungen zu Rollenkonzepten und Gestaltungsmethoden der MZKG sind deshalb TA-Ansätze von Interesse, die eine Identifikation und Berücksichtigung sozialer und ethischer Implikationen von KI-Systemen ermöglichen. Zur Ermittlung geeigneter Ansätze habe ich die TA-Ansatzlandschaft auf Gestaltungsansätze analysiert, die ein entsprechend ausreichendes Potenzial besitzen. Als geeignet erachte ich Ansätze, die anstelle einer rein passiven und abstandwahrenden Betrachtung von Gestaltungsprozessen

sen aktive Einflussnahmen in das Gestaltungsgeschehen vorsehen und dazu möglichst konkrete Methoden vorschlagen. Geeignete Ansätze müssen es Gestaltungsverantwortlichen ferner erlauben, eine Abwägung zwischen technischen, sozialen und ethischen Gestaltungsfaktoren zu ermöglichen, um begründbare Gestaltungsentscheidungen zu treffen. Diese Analyse lies drei Ansätze hervortreten: (1) das CTA, das danach strebt, bereits in frühen Phasen von Gestaltungsprozessen die Reflexion über gesellschaftliche Auswirkungen anzuregen; (2) das VSD, das danach strebt, die Auseinandersetzung mit ethischen Aspekten systematisch in Gestaltungsprozesse zu integrieren; und (3) das Responsible Research and Innovation (Burget et al., 2017; Owen et al., 2013; Von Schomberg, 2013), das danach strebt, die Folgen von Gestaltungsprozessen für Gesellschaft und Umwelt stärker in den Fokus zu rücken. Aufgrund der hohen Relevanz sozialer und ethischer Fragestellungen in HCAI widme ich mich in der Dissertation besonders der Betrachtung von CTA und VSD.

### 1.3 Struktur der Arbeit

Mit Fokus auf die beiden Ansätze CTA und VSD lautet die übergeordnete Forschungsfrage der Dissertation: Welche Erkenntnisse kann Human-centered AI (HCAI) von den Ansätzen Constructive Technology Assessment (CTA) und Value Sensitive Design (VSD) aufgreifen, um KI-Systeme menschenzentriert zu gestalten?

Zur Beantwortung dieser Frage fokussiere ich zunächst die in CTA und VSD verwendeten Rollenkonzepte. Die verwendeten Rollenkonzepte sind von zentralem Interesse, da in ihnen die in CTA und VSD zugrunde gelegten Annahmen zu menschlichem Leben und Handeln zum Ausdruck kommen, und die zu Gestaltungszwecken methodisch instrumentalisiert werden. Die Rollenkonzepte repräsentieren die verschiedenen Dimensionen menschlichen Lebens und Handelns, die in beiden Ansätzen in Bezug auf Gestaltungsprozesse für relevant erachtet werden, weshalb sie ein zentrales Kriterium darstellen, um die Interpretationen von Menschorientierung in CTA und VSD nachzuvollziehen. Ein Verständnis über die Auffassungen von Menschorientierung ermöglicht ferner, sie mit jenen in HCAI zu vergleichen und aus diesem Vergleich Rückschlüsse zur konzeptuellen Anschlussfähigkeit von CTA und VSD zu HCAI abzuleiten. Entsprechend lautet die erste Forschungsfrage der Dissertation: Welche Rollenkonzepte sind in CTA und VSD vorhanden und welche Zwecke sollen sie erfüllen?

Die häufig stark impliziten Bedeutungen der Rollenkonzepte machen es erforderlich, bei der Ergründung ihrer Bedeutungen und Zwecke auch auf relevante (inter-)disziplinäre und historische Hintergründe einzugehen, die mit den Rollenkonzepten verbunden sind. Die Analyse der Ursprünge der Rollenkonzepte ermöglicht die Freilegung zusätzlich relevanter Informationen zu den mit den Rollenkonzepten verbundenen konzeptuellen und methodischen Intentionen in CTA und VSD und vervollständigt damit die Beschreibung der Bedeutungen und Zwecke der Rollenkonzepte um wertvolle Erkenntnisse, die ohne eine Beleuchtung ihrer Ursprünge verborgen blieben. In Ergänzung zu Forschungsfrage 1 lautet Forschungsfrage 2 der Dissertation deshalb: Wie kommen die Rollenkonzepte zustande?

Eine weitere Eigenschaft von Rollenkonzepten ist, dass sie aufgrund ihres Modellcharakters bestimmte Aspekte menschlichen Lebens und Handelns fokussieren und dafür andere Aspekte ausblenden. Eine derartige Reduktion erfolgt häufig, um auf konzeptueller Ebene Gedanken zu entwickeln, zu deren Entwicklung ein Setzen von Gültigkeitsgrenzen zielführend ist. Es handelt sich dabei um einen Kompromiss, der aus Zweckgründen eingegangen wird, mit dem jedoch unweigerlich einhergeht, dass Aspekte, die im Kontext menschenorientierter Gestaltungsprozesse potenzielle Relevanz besitzen, nicht betrachtet werden. Durch die Beleuchtung dieser blinden Flecke der verwendeten Rollenkonzepte von CTA und VSD wird es mir möglich, die Defizite aufzuzeigen, die sie angesichts der Bedarfe der MZKG sowie in HCAI aufweisen. Entsprechend lautet die dritte Forschungsfrage der Dissertation: Was wird in den Rollenkonzepten nicht abgebildet?

Nach der Untersuchung der konzeptuellen Anschlussfähigkeit wende ich mich der Methodik von CTA und VSD zu. In HCAI herrscht ein Bedarf an Gestaltungsmethoden vor, die es ermöglichen sollen, gesicherte theoretische Erkenntnisse in der Gestaltungspraxis zu verankern, zu erproben und so zur Realisierung menschenzentrierter KI-Systeme beizutragen. Dies erfordert die Durchführung von Operationalisierungsanstrengungen in HCAI, deren Herausforderungen häufig darin liegen, gangbare Wege zu finden, um einerseits den Anspruch von Wissenschaftlichkeit zu wahren und andererseits den in der Gestaltungspraxis vorherrschenden Anforderungen und Bedingungen Rechnung zu tragen. Die Gestaltungsmethoden, die in CTA und VSD entwickelt und verwendet werden, stellen Vorschläge dar, dieser notwendigen Kompromissfindung zu begegnen. Die in ihnen verwandten Strategien zur Kompromissfindung lohnen daher einen genaueren Blick, da sie möglicherweise zielführende Hinweise für die Operationalisierungsbemühungen in HCAI bereitstellen können. Jedoch stammen viele der CTA- und VSD- Methoden aus dem Zeitraum der 1990er- und 2000er-Jahre und wurden nicht explizit für die Gestaltung von KI-Systemen entwickelt. Aus diesem Grund ist eine Evaluation der Methoden erforderlich, in der die gegenwärtigen Möglichkeiten und Anforderungen der praktischen KI-Gestaltung als Kriterien dienen. Dies ermöglicht, fundierte Aussagen über die Anschlussfähigkeit der Methoden von CTA und VSD für die Bestrebungen einer MZKG in HCAI zu treffen. Vor diesem Hintergrund lautet die vierte Forschungsfrage der Dissertation: Inwieweit eignen sich die Methoden von CTA und VSD für eine Erwägung in HCAI?

Bei einer Evaluation der Methoden anhand Kriterien der Gestaltungspraxis ist zu erwarten, dass neben Elementen der inhaltlichen und organisatorischen Eignung auch Erkenntnisse zum Vorschein treten, die die Defizite der Methoden deutlich werden lassen. Auch diese halten wertvolle Erkenntnisse für die Zielsetzungen in HCAI bereit. Für das Feld HCAI können die Defizite der CTA- und VSD-Methoden aufzeigen, welche Eigenschaften HCAI-Methoden besitzen müssen, um dem Anspruch von Praxistauglichkeit gerecht zu werden. Darüber hinaus stärkt eine Evaluation anhand von Kriterien der KI-Gestaltungspraxis die Robustheit der CTA- und VSD-Methoden, womit der in den Ansätzen verankerte Anspruch, auf die Gestaltung sämtlicher Technologien anwendbar zu sein, an einer konkreten Beispieltechnologie evaluiert wird. Analog zur Untersuchung der Rollenkonzepte widme ich mich bei der Untersuchung der CTA- und VSD-Methoden deshalb bewusst auch ihren Schwachstellen. Ferner erlaubt mir das Einbeziehen der Defizite der Methoden differenzierte Aussagen über die Anschlussfähigkeit der CTA- und VSD-Methoden zu den Zielsetzungen einer MZKG in HCAI zu treffen. Daher lautet die fünfte Forschungsfrage der Dissertation: Was können die Methoden von CTA und VSD für HCAI nicht bereitstellen?

Ein weiterer Effekt der differenzierten Auseinandersetzung mit den CTA- und VSD-Methoden liegt darin, dass neben Stärken und Schwächen auch Potenziale der Methoden zum Vorschein treten, die bezüglich ihrer Eignung für die KI-Gestaltungspraxis bislang nur wenig ausgeschöpft werden. Da die Methoden beider Ansätze auch in Kontexten der KI-Gestaltung erwägbare sein sollen, stellen die Erkenntnisse zu Potenzialen Möglichkeiten und Impulse dar, die künftig in CTA und VSD aufgegriffen werden können, um diesem Anspruch Rechnung zu tragen. Sie ermöglichen es, Anpassungsnotwendigkeiten sowie Weiterentwicklungsmöglichkeiten aufzuzeigen, die einer Steigerung der Relevanz im KI- und HCAI-Diskurs zuträglich sein können. Die sechste Forschungsfrage der Dissertation lautet deshalb: Welche Anpassungen sind erforderlich, um die Praxistauglichkeit der Methoden von CTA und VSD zu fördern?

Zur Beantwortung dieser Fragen führe ich eine vergleichende Analyse der Ansätze CTA und VSD auf Basis einschlägiger Literatur durch, in Kombination mit der Entwicklung eines Kategoriensystems, das für die Forschungsfragen relevante Vergleichskategorien beinhaltet. Dies ermöglicht eine strukturierte Analyse der in den Ansätzen vorhandenen Rollenkonzepte sowie eine nachvollziehbare Darstellung der in CTA und VSD vorgeschlagenen Gestaltungsmethoden. Um die Praxistauglichkeit der Gestaltungsmethoden im KI-Kontext zu evaluieren führe ich eine qualitative Interviewstudie mit Experten der KI-Gestaltungspraxis durch. Zur inhaltlichen Unterstützung der Experteninterviews entwickle ich Storyboards, die die Ziele und Vorgaben der Gestaltungsmethoden in Form von zusammenhängenden Text- und Bildinformationen wiedergeben. Zur Auswertung der von den KI-Experten geäußerten Einschätzungen der CTA- und VSD-Methoden führe ich eine SWOT-Analyse durch (siehe Abb. 1).





Abbildung 1: Forschungsdesign der Dissertation

Zu Beginn der Dissertation stelle ich in Kapitel 2 die Grundlagen der Gestaltungsphilosophie der MZG vor. Ich gehe auf die ihr zugrundeliegenden Gedanken zu einer an menschlichen Eigenschaften und Bedürfnissen ausgerichteten IT-Gestaltung ein, erläutere die Gestaltungsprinzipien der MZG und gehe dabei auf relevante terminologische und inhaltliche Besonderheiten ein, die die Anwendung und Rezeption des Ansatzes auszeichnen. In Kapitel 3 schlage ich den Bogen zum Feld HCAI, in dem die Ideen und Gestaltungsprinzipien der MZG versucht werden auf die Gestaltung von KI-Systemen zu übertragen. Nach einer kurzen Skizze der aktuellen Forschungslandschaft des Feldes zeige ich, welche Ideen und Vorschläge zur Weiterentwicklung des MZG-Ansatzes in HCAI derzeit diskutiert werden und weshalb die damit verbundenen komplexeren Implikationen den derzeitigen Diskurs vor verschiedene Herausforderungen stellen. In Kapitel 4 führe ich einen Perspektivwechsel zu TA durch und stelle die zentralen Annahmen und Methoden in CTA und VSD vor. Dies ermöglicht mir, in einer Zwischenbetrachtung Charakteristika von TA und HCAI zu vergleichen und vor dem Hintergrund der übergeordneten Forschungsfrage zu diskutieren. Kapitel 5 ist der Untersuchung der Rollenkonzepte in CTA und VSD gewidmet. Nach einer Darlegung meines Verständnisses von Rollenkonzepten stelle ich nacheinander die Bedeutungen, Zwecke, Ursprünge und blinden Flecke der Rollenkonzepte vor und diskutiere die verschiedenen Schlussfolgerungen und Antworten, die für die Forschungsfrage der Dissertation Relevanz besitzen. Kapitel 6 beinhaltet sämtliche Erkenntnisse, die die Expertenevaluation der CTA- und VSD-Methoden hervorgebracht haben. Die Struktur des Kapitels ergibt sich aus der Durchführung der SWOT-Analyse, wodurch es mir möglich wird, die von den KI-Experten identifizierten Stärken, Schwächen, Chancen und Risiken der Methoden für die KI-Gestaltung konzise vorzustellen. Dies ermöglicht mir, die Beitragsfähigkeit der Methoden zu den Operationalisierungsbestrebungen in HCAI zu diskutieren und dabei auch Empfehlungen zur Weiterentwicklung im Sinne einer Praxisauglichkeit zu formulieren. Im abschließenden Kapitel 7 reflektiere ich zusammenfassend die wesentlichen Elemente, die eine Anschlussfähigkeit zwischen CTA, VSD und HCAI bedingen.

## 2 Menschzentrierte Gestaltung

Den Ausgangspunkt meiner Betrachtungen stellen die Ideen, Konzepte und Vorschläge dar, die im Gestaltungsansatz der MZG zum Tragen kommen. In diesem Kapitel stelle ich die in ihm verankerten Gedanken vor, die seit den 1970er-Jahren von verschiedenen Vordenkerinnen und Vordenkern geprägt wurden (siehe Kap. 2.2). Ich gehe zudem auf die Gestaltungsprinzipien der MZG ein, für die die Norm DIN EN ISO 9241-210 sowohl im akademischen Diskurs wie auch in der Gestaltungspraxis eine zentrale Ressource darstellt (siehe Kap. 2.3). Aus der Auseinandersetzung mit diesen Aspekten wird ersichtlich, wie in der MZG danach gestrebt wird, technikwissenschaftliche, sozialwissenschaftliche und in Ansätzen auch geisteswissenschaftliche Erkenntnisse zu verknüpfen, um eine erfolgversprechende menschenzentrierte Gestaltung von IT-Systemen zu ermöglichen. Auch wird es mir dadurch möglich, im anschließenden Kapitel 3 die im Feld HCAI vorhandenen Denkmuster und die Betrachtungsweisen der Herausforderungen zur MZKG einzuordnen.

### 2.1 Einführung

Das Verhältnis von Mensch und Technik wird seit dem 20. Jahrhundert stark durch die Informationstechnologie geprägt. Sie hat grundlegend verändert, wie Menschen arbeiten, lernen, denken und kommunizieren (Cox & Marshall, 2007; Karaman Aksentijević et al., 2021; Saldanha et al., 2021). Entscheidenden Einfluss auf solche Entwicklungen haben Gestalter, die in der IT-Gestaltung konzipierend, entwickelnd, organisierend oder auf andere Weise verantwortungstragend tätig sind und als Urheber von Soft- und Hardware-systemen die Möglichkeiten zur Aneignung und Nutzung der Systeme maßgeblich formen. In den Feldern der IT-Gestaltung hat sich dabei sukzessive die Erkenntnis durchgesetzt, dass mit gestalterischem Handeln auch eine Verantwortung für kulturelle, soziale und ethische Fragestellungen und Auswirkungen verbunden ist, weshalb die Auseinandersetzung mit diesen Aspekten stärker in den Fokus der Diskurse zur IT-Gestaltung gerückt ist. Verschiedene inter- und transdisziplinäre Communities wie beispielsweise die HCI (Carroll, 1997; Helander, 2014; I. S. MacKenzie, 2024; B. Myers et al., 1996; B. A. Myers, 1998), das Computer Supported Cooperative Work (CSCW; Borghoff & Schlichter, 2000; Gross & Koch, 2007; Grudin, 1994; Palmer & Fields, 1994), oder die Informationsethik (Bendel, 2019; Capurro, 2004; Heesen, 2021; Lenzen, 2023) sind inzwischen etablierte Bestandteile der Debatten zur IT-Gestaltung und führen Wissen aus Feldern wie der Ergonomie, Psychologie, Soziologie oder Anthropologie zusammen. Neben dem Verknüpfen von Wissen ist ein weiterer Schwerpunkt des interdisziplinären Diskurses die Operationalisierung von Wissen für die Gestaltungspraxis, was in der Vergangenheit Gestaltungsansätze wie das Participatory Design (PD; Bannon & Ehn, 2012; Björgvinsson et al., 2010; Kensing & Blomberg, 1998; Spinuzzi, 2005), das SCRUM (Hron & Obwegeser, 2022; Permana, 2015; Schwaber, 1997; Srivastava et al., 2017) oder das Design Thinking (Brown, 2008; Dorst, 2011; Liedtka, 2018; Razzouk & Shute, 2012) hervorgebracht hat. Auch der aus dem agilen Manifest (Beck et al., 2001) hervorgegangene Ansatz Agile zählt

zu diesen Ansätzen (Eckstein, 2013; Koch, 2004; Shore & Warden, 2021) und muss aufgrund seiner Relevanz für die Gestaltungspraxis Erwähnung finden. Solche Ansätze ermöglichen Gestaltern und Verantwortungsträgern eine deutlich gesteigerte Mensch- und Kontextorientierung in Gestaltungsprozessen, da sie mehr Raum und mehr Ressourcen für die Auseinandersetzung mit relevanten Stakeholdern vorsehen, und durch ihre mehr zyklische Vorgehensweise die in der Gestaltungspraxis benötigte Flexibilität erhöhen.

Der Ansatz der MZG (z.B. Heinecke, 2012; Nelles et al., 2018; Povse, 2023), im Englischen Human-centered Design (HCD; Boy, 2011; Cooley, 2000; Gasson, 2003; Oviatt, 2006; Winograd & Woods, 1997) ist ein weiteres Resultat der interdisziplinären Auseinandersetzung mit Fragen der IT-Gestaltung, der die Gestaltungstheorie und die Gestaltungspraxis seit den 1990er-Jahren entscheidend geprägt hat. Am Namen des Ansatzes wird bereits der deutliche Anspruch der Menschorientierung ersichtlich, der sich unmissverständlich entgegen einer technikzentrierten Gestaltung von IT-Systemen positioniert, die primär auf die technischen Möglichkeiten und Unmöglichkeiten fokussiert ist (z.B. Norman, 1994). Ein weiteres Merkmal des Ansatzes der MZG ist die Verbindung sozialwissenschaftlicher Erkenntnisse mit technisch fundiertem, aus der Gestaltungspraxis gespeistem Erfahrungswissen. Dieser zweite Aspekt ist von besonderer Bedeutung, da er zeigt, dass die MZG kein Produkt eines ausschließlich von außen an ihre angestammte Disziplin der Informatik herangetragenen Anliegens ist, sondern sich auf Anteile einer genuin intrinsisch motivierten Strömung der Informatik gründet. Die Entwicklung des Ansatzes der MZG ist dadurch von Beginn an durch Bemühungen gekennzeichnet, relevante Erkenntnisse und Beobachtungen aus der Gestaltungspraxis mit jenen der Gestaltungstheorie zu verknüpfen. Ein Ergebnis dieser Bemühungen ist das in der Norm DIN EN ISO 9241-210 (DIN EN ISO 9241-210:2020-03, 2020) vorgeschlagene Modell zur Realisierung menschenzentrierter Gestaltungsprozesse (siehe Abb. 2)

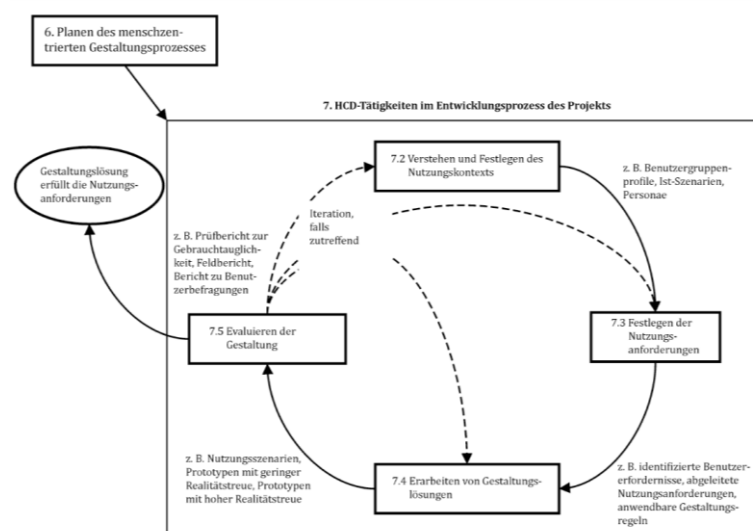


Abbildung 2: Ein menschenzentrierter Gestaltungsprozess nach DIN EN ISO 9241-210:2020-03 (2020, S. 25)

Die Vorteile und Fortschritte dieses zyklischen Modells werden am deutlichsten, wenn sie dem linearen Wasserfallmodell gegenübergestellt werden (z.B. Adenowo & Adenowo, 2013; Boehm, 2006; Raccoon, 1997), das bis in die 1990er-Jahre einen Standard in der Entwicklung von Softwaresystemen darstellte (siehe Abb. 3).

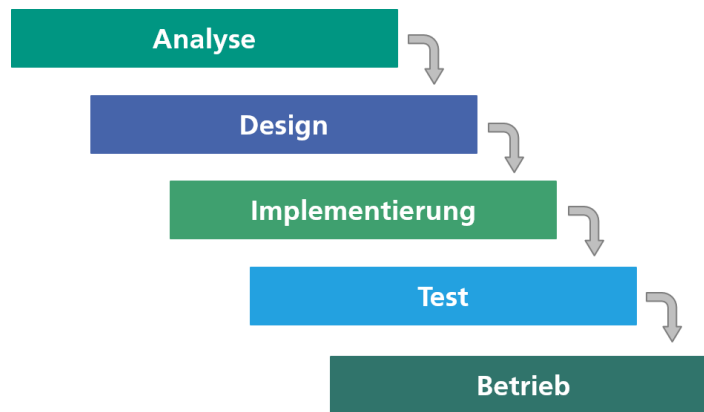


Abbildung 3: Das Wasserfallmodell

Die wesentlichen Schwächen des Wasserfallmodells liegen in seiner linearen Vorgehensweise, wodurch es anfällig ist für die Durchführung notwendiger Korrekturmaßnahmen im Gestaltungsgeschehen, sowie der Adressierung von Anforderungsveränderungen, die im Laufe von Gestaltungsprozessen auftreten können. Resultate von Gestaltungsprozessen, die nach dem Wasserfallmodell strukturiert werden, sind aufgrund einer mangelhaften Fehlerkultur sehr kostspielig und haben häufig Systeme zum Ergebnis, die bereits einfachste Usabilitykriterien nicht erfüllen können. Diese Schwächen werden in zyklischen Vorgehensmodellen wie jenem der MZG behoben, doch sind auch solche Modelle defizitär. Ein Kritikpunkt der auf solche Modelle angebracht werden kann ist, dass durch eine zu starke Fokussierung auf den Menschen der Fokus auf die Aktivitäten verloren gehen kann, und dadurch der initiale Zweck der Gestaltungsbemühungen verfehlt wird (Norman, 2005). Eine weitere potenzielle Schwäche betrifft das Element der Partizipation von Nutzern, die häufig als Bestandteil menschzentrierter Gestaltungsprozesse vorgesehen sind – eine solche sollte nicht aus reinem Selbstzweck oder aufgrund (organisations-)politischer Motive erfolgen, sondern Ergebnis einer sorgfältigen Abwägung der Verantwortungsträger hinsichtlich einer hohen Aussicht auf tatsächlichem Mehrwertsgewinn für die Gestaltungsziele sein (Steen, 2012; van Velsen et al., 2022). Und auch in Gestaltungskontexten mit tendenziell hohen soziotechnischen Komplexitätsgraden, wie sie häufig in der KI-Gestaltung vorzufinden sind, stößt das Vorgehensmodell an seine Grenzen (siehe Kap. 3). Damit sind, stark verkürzt, die wesentlichen Elemente der MZG dargestellt. Die Ursprünge und die Gestaltungsprinzipien der MZG lohnen hingegen einen detaillierteren Blick, da beide Elemente auch in den HCAI-Debatten um eine MZKG von zentraler Bedeutung sind.

## 2.2 Historie und Grundgedanken

Die Bemühungen zur Ausarbeitung und Weiterentwicklung der MZG zeugen von großem Engagement, haben jedoch auch dazu geführt, dass sich um die MZG eine große terminologische Heterogenität herausgebildet hat. So werden in der Literatur und in der Gestaltungspraxis verschiedene Inhalte und Konzepte mit der MZG assoziiert, wenn zusätzlich zum Adjektiv menschenzentriert von einer menschenorientierten (Cho & Lee, 2002; Saleem et al., 2015; Stern & Becker, 2019), menschengerechten (Coy et al., 2013; Fischer, 1986), (be-)nutzerzentrierten (Abrams et al., 2004; Mädche & Wallach, 2014; Mao et al., 2005), oder auch sozialverträglichen Gestaltung (Floyd, 1987; D. Lee et al., 2018; von Alemann et al., 1986) von IT-Systemen gesprochen wird. Diese Vielfalt zeigt zweierlei: zum einen sind im Feld der MZG verschiedene disziplinäre Färbungen erkennbar, da manche der Adjektive beispielsweise mehr in Feldern wie der Ergonomie zum üblichen Sprachgebrauch zählen. Zum anderen verhalten sich die unter diesen Labels dargelegten Gedankengänge auf inhaltlicher Ebene häufig komplementär zueinander, wovon die Ausarbeitung der MZG in der Vergangenheit in mehrerlei Hinsicht profitieren konnte. Dennoch scheint es angesichts dieser Sachlage hilfreich zu sein, jene historischen Elemente und Grundgedanken zu beleuchten, die für die Fragestellung der Dissertation Relevanz besitzen. Instrukтив sind dafür die Arbeiten von Joseph Weizenbaum, Christiane Floyd, Donald Norman und Lucy Suchman, da sie bereits früh auf entscheidende Weise zur Entwicklung und Etablierung der MZG beigetragen haben. Durch unterschiedliche Schwerpunktsetzungen legen sie die Notwendigkeiten und Möglichkeiten einer MZG von IT-Systemen dar und zeigen punktuell auch die Verbindungslinien zur KI-Gestaltung auf.

### 2.2.1 Kritik unreflektierter Computereuphorie (Joseph Weizenbaum)

Indem er unermüdlich die Besonderheiten und Qualitäten menschlichen Lebens denen von IT-Systemen gegenüberstellt und die daraus resultierenden Unterschiede in den Kontext von Gestaltungsprozessen einordnet, bereitet Joseph Weizenbaum früh den Weg für die Entwicklung für Ansätze wie die MZG. Er gilt heute u.a. deshalb als „Kritiker unreflektierter Computereuphorie“ (Wendt & Klug, 2001), da er offen die seiner Ansicht nach dehumanisierenden und technikzentrierten Tendenzen der Informatik seiner Zeit kritisierte. Als Professor für Computer Science am Massachusetts Institute of Technology stellt er in den 1960er- und 1970er-Jahren zunächst viele Weichen in der Informatik, unter anderem durch seine Mitwirkung an der Entwicklung des Vorläufersystems des heutigen Internets („ARPAnet“; Abbate, 2003; Roberts & Kahn, 1972). Als großer Meilenstein für die KI-Forschung und gleichzeitig Wendepunkt seiner Laufbahn gilt das von ihm entwickelte Programm ELIZA (Weizenbaum, 1966, 1967). ELIZA war in der Lage, ein psychotherapeutisches Gespräch auf Basis natürlicher Sprache zu simulieren, indem es bestimmte Schlüsselwörter aus den von Menschen getätigten Inputs in seinen Rückfragen aufgriff. Das Programm wurde sowohl in der Öffentlichkeit wie auch der Fachwelt eifrig diskutiert, da es die Fähigkeiten eines vermeintlich künstlich intelligenten Systems hinsichtlich der Verarbeitung natürlicher Sprache auf bislang ungekannte Weise demonstrierte. Die fachliche

und öffentliche Resonanz zu seinem System war für Joseph Weizenbaum selbst jedoch Anlass zu Sorge. Wie sehr die Menschen ELIZA anthropomorphisieren („ELIZA-Effekt“; z.B. Kim et al., 2019; Wilson, 2010) und wie sehr ELIZA auch Hoffnungen schürt, psychotherapeutische Gespräche und Therapiesitzungen künftig stärker zu automatisieren, prägen ihn nachhaltig:

*„My own shock was administered [...] by some people who insisted on misinterpreting a piece of work I had done“* (Weizenbaum, 1976, S. 2).

Er sieht sich daraufhin dazu veranlasst, sich mit den philosophischen und gesellschaftlichen Zusammenhängen seiner eigenen Arbeit und auch der seines angestammten Fachbereichs der Informatik zu befassen. Dies führt dazu, dass er sich im weiteren Verlauf seines Lebens kritisch zu vielen Entwicklungen v.a. auch in der KI-Forschung positioniert und auch die Öffentlichkeit für ihre Technikgläubigkeit kritisiert. Die Kritik eben jenes Fachgebiets, zu welchem er selbst zuvor viele entscheidende Beiträge erarbeitet hatte, bringt ihm den Ruf eines „Nestbeschmutzers“ ein (Kirkamm, 1995), und auch er selbst bezeichnet sich an einer Stelle als einen „Dissidenten der Informatik“ (Weizenbaum, 1991).

Seinem Richtungswechsel verleiht Joseph Weizenbaum mit dem Buch *Computer Power and Human Reason – From Judgement to Calculation* (Weizenbaum, 1976) Konturen. Darin verarbeitet er den seiner Ansicht nach grundlegenden Unterschied zwischen Mensch und Computer, den er auch schon im Titel hervorhebt – Menschen sind nach seiner Auffassung in der Lage zu urteilen (judgement), besitzen also die Fähigkeit, eine Wahl zu treffen, die ihre Werte widerspiegelt und in die ebenso Emotionen, Empathie und Lebenserfahrung miteinfließen. Computer hingegen besäßen keine solcher menschlichen Qualitäten, da sie lediglich in der Lage seien, Berechnungen (calculations) anzustellen. Diese Unterscheidung macht er insbesondere für KI geltend, die er unmissverständlich von menschlicher Intelligenz unterscheidet (Weizenbaum, 1976, S. 203). Seine Unterscheidung zwischen Mensch und Computer ficht die zur Zeit der späten 1970er-Jahre in der Informatik weit vertretene Annahme an, die Instrumente und Erzeugnisse von IT-Gestaltungsprozessen seien wertfrei. Joseph Weizenbaum fordert insbesondere die Gestalter von IT-Systemen dazu auf, sich mit den Besonderheiten menschlichen Lebens auseinanderzusetzen und diesen mehr Aufmerksamkeit bei ihrer Arbeit zu widmen. Besonders die zeitgenössische KI-Forschung kritisiert er für ihre seiner Ansicht nach dehumanisierenden Tendenzen und Einstellungen (Wendt & Klug, 2001, S. 39). Damit konfrontiert er die Verantwortungsträger in den Feldern der IT- und KI-Gestaltung offen mit ihrer Verantwortung, insbesondere gegenüber ihren technisch weniger versierten Mitmenschen und zeigt auch in dieser Beziehung die bislang wenig beachtete Verbindung zwischen gestalterischem Handeln und ethischer Verantwortung auf. Jedoch sieht er nicht nur die Gestalter in der Verantwortung für ein wünschenswertes Verhältnis zwischen Menschen und IT-Systemen, auch der Gesellschaft und dem Individuum trägt er auf, im öffentlichen Diskurs die Position des Computers in der Gesellschaft in Frage zu stellen und sie für das gesellschaftliche Miteinander aktiv auszuhandeln. Sein Aufruf an die Öffentlichkeit, die von Gestaltern stammenden Versprechungen kritisch zu hinterfragen und diesen gegenüber keine Gleichgültigkeit walten zu lassen, prägt auch seine Arbeit späterer Jahre. In diesen

bringt er auch besonders seine Befürchtung zum Ausdruck, dass Menschen aufgrund von Computern zunehmend ihre Autonomie und Freiheit verlieren oder gar unwissend aufgeben. In seinem gesamten Schaffen wird die Bedeutung von Selbstbestimmung und Mündigkeit der Menschen deutlich, die er auch als Argument dafür verwendet, weshalb die Gestaltung von IT-Systemen im übergeordneten Zusammenhang nicht von deren Einbettung in soziale Kontexte getrennt betrachtet werden kann.

Manche der Positionen Joseph Weizenbaums werden mittlerweile stark kritisiert (z.B. Sarasin, 2022), auch da sie aufgrund der teilweise großen technischen Fortschritte der Informatik an Relevanz verloren haben. Auch wenn solche Kritiken berechtigt sind, haben Joseph Weizenbaums Bemühungen zur Entmystifizierung von IT-Systemen und deren Funktionsweise auf technischer Ebene eine fachliche und öffentliche Diskussion angestoßen. Diese waren entscheidend für die Ausbildung reflektierter Mindsets, die zur Entwicklung und Anwendung von Gestaltungsansätzen wie der MZG und ihrer deutlich höheren Orientierung an den Eigenschaften menschlichen Lebens notwendig sind.

### 2.2.2 Integratives Verständnis von Gestaltung und Nutzung (Christiane Floyd)

Christiane Floyd hat durch ihre Auseinandersetzung mit Methoden des Software Engineerings (SE) besonders zur Klärung der methodischen Frage der MZG beigetragen. Als Professorin für Informatik an der Technischen Universität Berlin und später an der Universität Hamburg setzt sie sich dafür ein, dass Softwaregestaltung und -nutzung weniger als rational getrennte Elemente betrachtet werden, sondern mehr als wechselseitiger und kommunikativer Lernprozess zwischen Gestaltern und Nutzern (Floyd, 1988). Die traditionelle Trennung beider Seiten stellt für sie eine Kluft dar (Wolff, 2021), auf die sie die unzureichende methodische Ausgangslage für eine menschenorientierte Gestaltung im SE zurückführt (Floyd, Mehl, et al., 1989, S. 274). Ihr Beitrag zur Überbrückung dieser Kluft ist der Ansatz „Softwaretechnik für evolutionäre und partizipative Systementwicklung“, kurz STEPS (Floyd, 1987; Floyd, Reisin, et al., 1989). STEPS liegt ein „menschenzentrierter Qualitätsbegriff“ (Floyd, 1987, S. 500) zugrunde, der primär nach dem intendierten Einsatz einer Software fragt und die Gestaltungsaktivitäten an diesem orientiert. Mit dieser Vorgehensweise wird eine Benutzerfreundlichkeit angestrebt, die deutlich über den Charakter eines Lippenbekenntnisses hinausgeht (Wolff, 2021) und adressiert zudem die identifizierte Notwendigkeit, sich im SE neben den formalisierbaren auch mit den informellen Gesichtspunkten von Gestaltungstätigkeiten auseinanderzusetzen. In STEPS kommt zudem ein Verständnis von Design zum Tragen, welches sich durch „das Ringen um Qualität“ (Floyd, 1989, S. 3) auszeichnet. Dieses Qualitätsverständnis erstreckt sich dabei auf das äußerliche Erscheinungsbild eines Systems, auf dessen technische Funktionalität sowie auf seine Möglichkeiten zur Einbettung in menschliche Handlungskontexte.

Besonders der Blick auf die Einbettungsmöglichkeiten soll die soziotechnischen und kulturellen Aspekte eines Gestaltungsprozesses stärker in den Fokus rücken und ist inspiriert von den damaligen SE-Ansätzen skandinavischer Länder (siehe Abb. 4).

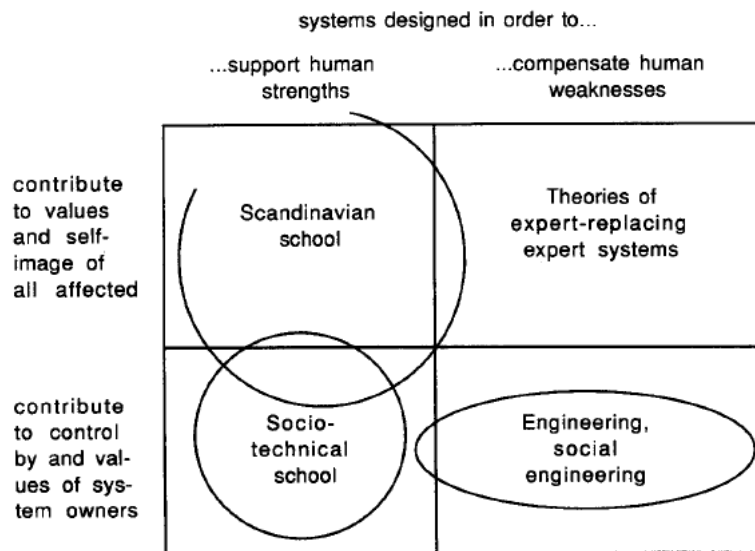


Abbildung 4: Darstellung des skandinavischen Ansatzes zur Softwareentwicklung gegen Ende der 1980er-Jahre (Floyd, Mehl, et al., 1989, S. 265)

Mit diesen beschäftigt sich Christiane Floyd intensiv und leitet daraus ihr Verständnis für erstrebenswertes Softwaredesign ab. Dieses strebt danach, die vorgefundenen und teilweise widersprüchlichen sozialen und technischen Belange in Einklang zu bringen und dazu die Zusammenarbeit zwischen Wissenschaft, Praxis und Gesellschaft stärker zu fördern. Dieses erweiterte Verständnis von SE wird in das methodische Konzept von STEPS transferiert (siehe Abb. 5).



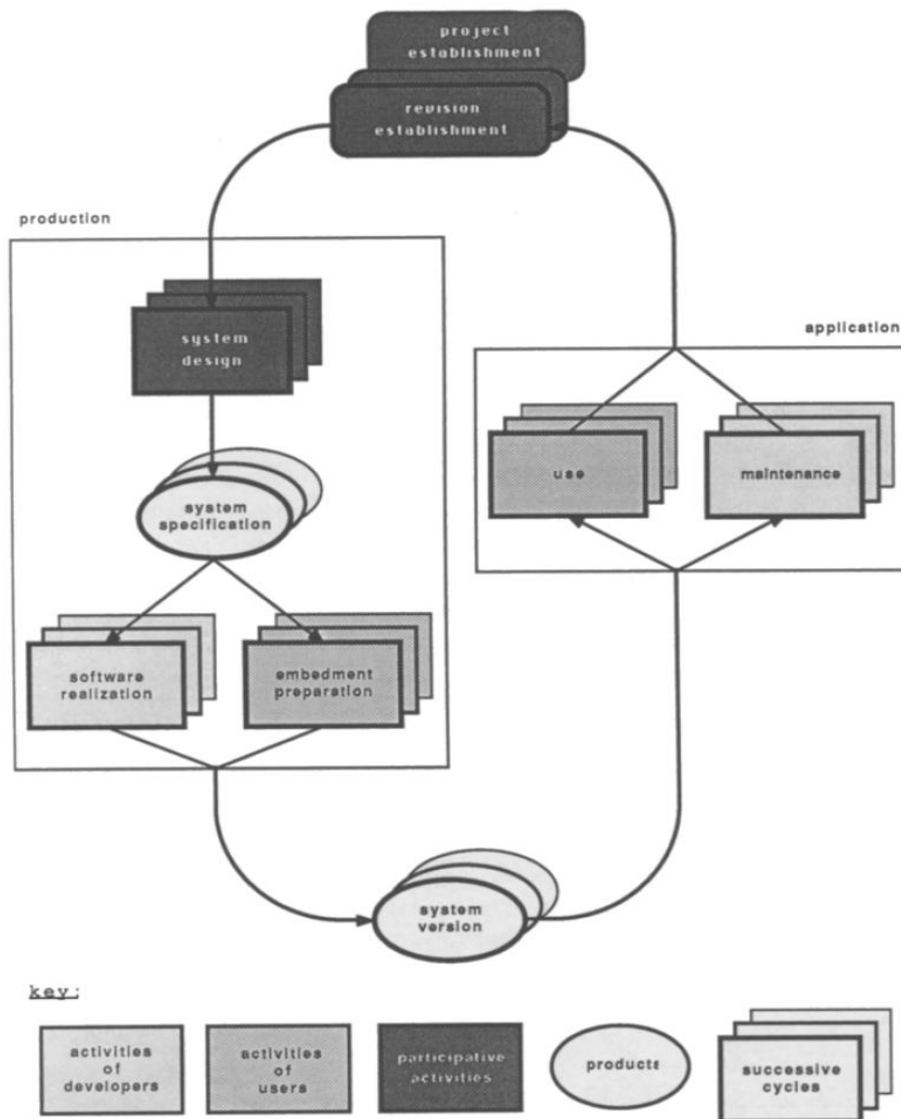


Abbildung 5: Prozessmodell von STEPS (Floyd, Reisin, et al., 1989, S. 58)

Ein grundsätzliches Merkmal der STEPS-Methode ist die zyklische Ausrichtung des SE-Prozesses. Speziell in der Wiederaufnahme bzw. kontinuierlichen Fortsetzung von Gestaltungsaktivitäten („production“-Phase im Modell) nach Erprobungsphasen im tatsächlichen Nutzungskontext („application“-Phase im Modell) sieht Christiane Floyd eine Chance, die Anpassung von Software an tatsächliche Nutzerbedürfnisse und Kontextgegebenheiten zu fördern. Dadurch rückt STEPS ab von der Vorstellung einer idealen und linearen Gestaltungstrategie, die Ende der 1980er-Jahre im SE noch weitgehend üblich war (Wasserfallmodell). Die zyklische Gestaltungsmethode bietet zudem die Chance, die Planung der für ein Projekt zur Verfügung stehenden finanziellen und personellen Ressourcen flexibler zu gestalten, wodurch Spielräume für die Adressierung sich als besonders wichtig herausstellender Aspekte eröffnen. Ein weiteres Novum der Methode von STEPS ist die Zusammen-

führung der Tätigkeiten von Gestaltern wie auch Nutzern zu bestimmten Zeitpunkten bzw. für bestimmte Aktivitäten im Prozess („participative activities“ im Prozess-Modell). Beide Seiten erarbeiten zusammen zu Beginn eines Projekts den übergeordneten Zweck der Software, führen gemeinsam Revisionen von Softwareversionen durch und transferieren die Revisionserkenntnisse kooperativ in die darauffolgende Systemversion.

STEPS und die darin operationalisierten Ideen zu Prototyping, Design und Partizipation (Floyd, 1984) sind wegweisend für die darauffolgenden Jahrzehnte, in denen die Bestrebungen zur Förderung partizipativer und menschenorientierter Elemente im SE fortgesetzt werden (Blomberg & Henderson, 1990; Bratteteig et al., 2012; Pipek & Wulf, 2009; Wulf & Rohde, 1995). Christiane Floyd setzt die Weiterentwicklung des Ansatzes dahingehend fort, dass sie ihm Erkenntnisse der Arbeitswissenschaft zuführt (Floyd & Piepenburg, 1993) und ihn in der Praxis erprobt (Floyd et al., 1997). Die Förderung antizipierender Tätigkeiten, die stärkere Orientierung an tatsächlichen Arbeitsvorgängen und die Verbindung von Kreativität und geplantem Vorgehen tragen sukzessive zu einem Paradigmenwechsel im SE bei, die instrumentell für Ansätze wie die MZG waren. Auch die Arbeit Joseph Weizenbaums ist in der Arbeit Christiane Floyds erkennbar (beide waren befreundet), da sie sich kontinuierlich und kritisch mit den sozialen Fragestellungen im SE auseinandersetzt (Dittrich et al., 2002; Floyd, 1992, 2005) und diese punktuell auch auf KI überträgt (Floyd, 1994). Anfang 2023 erläutert sie anlässlich des 100. Geburtstages Joseph Weizenbaums dessen Einfluss auf ihre Arbeit (Floyd, 2023) und stellt ihren Standpunkt zur menschenzentrierten IT-Gestaltung vor (siehe Abb. 6). Darin hebt sie hervor, dass man sich im SE von der Vorstellung verabschiedet hat, eine Systemanalyse könne objektiv sein. Ferner hat sich im SE die Erkenntnis durchgesetzt, dass nicht alles (z.B. eine Gesellschaft) gestaltbar sei, sondern dass das SE nur punktuell intervenieren könne. Der Trend in der MZG kehre sich ab von einer einmaligen Problemlösung, und wandle sich hin zu einer Prozessorientierung.

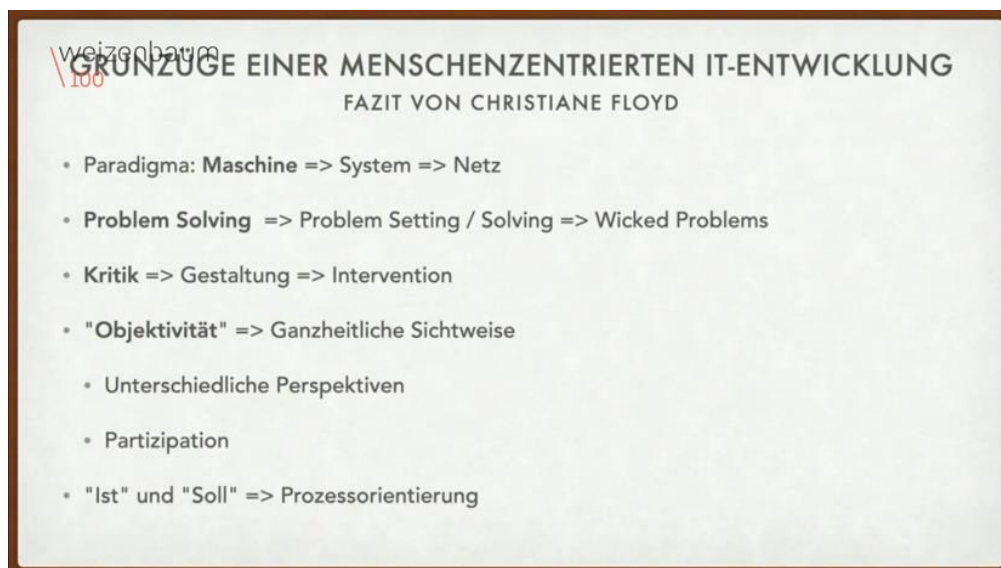


Abbildung 6: Christiane Floyds Position zur MZG im Jahr 2023 (Floyd, 2023)

### 2.2.3 Anbindung von Erkenntnissen der Psychologie und der Designforschung (Donald Norman)

In der MZG werden viele Erkenntnisse aus der Psychologie und der Designforschung auf die Gestaltung von IT-Systemen angewendet. Maßgeblich beteiligt an der Zuführung solchen Wissens ist Donald Norman in seiner Funktion als Professor für Cognitive Science & Psychology an der University of California und später als Professor für Design an der Northwestern University. Seine Arbeit ist zudem geprägt durch die kontinuierliche Verknüpfung von Forschung und Praxis, beispielsweise durch Kooperationen mit der NASA und Apple, durch die von ihm mitgegründete Nielsen Norman Group, oder auch durch seine Funktion im Executive Board der Interaction Design Foundation. Ausgehend von seinen Untersuchungen zum menschlichen Erinnerungsvermögen und kognitiver Informationsverarbeitung (z.B. Lindsay & Norman, 1972; Norman, 1982; Norman & Rumelhart, 1981) wendet er sich zu Beginn der 1980er-Jahre IT-Systemen zu. Er strebt an zu ergründen, inwieweit IT-Systeme kognitive Vorgänge des menschlichen Gehirns in ihren Interaktionskonzepten aufgreifen und berücksichtigen. Seine Untersuchungen zeigen große Defizite, was beispielsweise am Unix Texteditor „Ed“ deutlich wird. Dieser verfügt über keine automatische Speicherfunktion für Texteingaben, ebenso wenig wie über einen Dialog, der Nutzer an das manuelle Speichern erinnert (Norman, 1981). Solche Einblicke motivieren ihn zur Fortsetzung und Intensivierung der Auseinandersetzung mit IT-Systemen, wobei Donald Norman sich besonders mit der Gestaltung von User Interfaces (UIs) auseinandersetzt. Er macht sich zur Aufgabe, die Bedeutung von UIs von einem mehr zufälligen, wenig beachteten Nebenprodukt hin zu einem bewusst gestalteten Element in IT-Gestaltungsprozessen zu fördern. Seine Hauptkritik an den IT-Systemen der 1980er- und 1990er-Jahre ist die für kognitive Prozesse des Menschen zu hohe Komplexität, die im Ergebnis zu Verwirrung, Frust und schlussendlich Fehlern in Interaktionsvorgängen führen können. Besonders letztere stellen für ihn Kennzeichen mangelhafter Designvorgänge dar (Norman, 2013) und verdeutlichen für ihn die Notwendigkeit, den Bedürfnissen und Intentionen von Nutzern in Gestaltungsprozessen deutlich mehr Aufmerksamkeit zu widmen. Um seine Argumente zu verdeutlichen, zergliedert er Anforderungen und Handlungsintentionen, die von Menschen an IT-Systeme gestellt werden, um aufzuzeigen, welche Teilaspekte dieser Anforderungen in Gestaltungsprozessen wenig Beachtung finden oder gänzlich ignoriert werden (siehe exemplarisch Abb. 7).

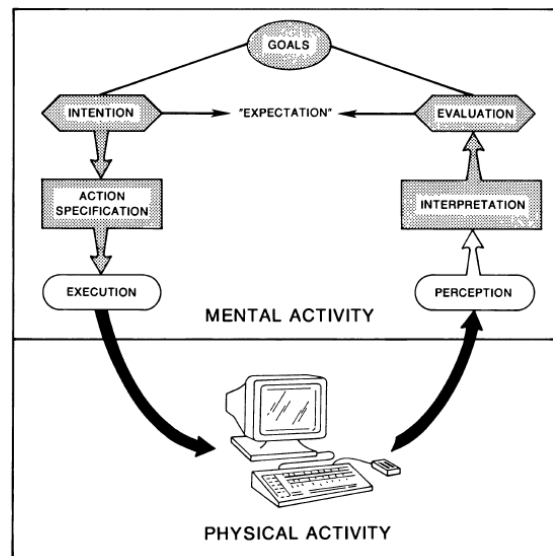


Abbildung 7: Elemente von Interaktionstätigkeiten mit Computersystemen (Norman, 1986, S. 42)

Sein Vorschlag für einen entsprechenden Gestaltungsansatz versteht er 1986 zunächst mit dem Namen User-centered Systems Design, kurz auch User-centered Design (UCD; Norman & Draper, 1986). Dieser Ansatz sieht vier Gestaltungsprinzipien vor: (1) einen eigenen Forschungszweig für UCD zu etablieren, in dem (2) die Gestaltung von UIs als zentrales Element von IT-Systemen erachtet wird. Darüber hinaus soll (3) die Gestaltung von UIs von der Gestaltung anderer Systemkomponenten separiert werden und (4) die Bedürfnisse von Nutzern als maßgeblichen Ausgangspunkt sämtlicher Aktivitäten der IT-Systemgestaltung erachtet werden. Bei diesen Gestaltungsprinzipien werden Parallelen zum (etwa zeitgleich vorgestellten) STEPS-Modell von Christiane Floyd erkennbar, jedoch argumentiert Donald Norman aus der Position eines Kognitionspsychologen. Er setzt die Weiterentwicklung des UCD-Ansatzes in den 1990er- und 2000er-Jahren fort, was neben Veränderungen auf sprachlicher Ebene auch Veränderungen auf inhaltlicher Ebene zum Ergebnis hat. So wandelt sich das UCD zum HCD und stellenweise auch zum People-centered Design (Interaction Design Foundation, 2024), wobei sich Donald Norman auf seine gewonnene Erkenntnis beruft, dass die Konzeption von Menschen als Nutzer Limitierungen unterliegt und sogar dehumanisierende Tendenzen aufweisen kann (Norman, 2006). Den zu Beginn der 2000er-Jahre als HCD bezeichneten Gestaltungsansatz charakterisieren jetzt die vier Gestaltungsprinzipien (1) der Ermittlung grundlegender Probleme von Menschen (nicht nur die Problemsymptome) sowie (2) explizit Menschen (und nicht Technik) in den Fokus zu rücken. Dazu sollen Gestalter und Entscheidungsträger (3) die relevanten Gestaltungsfaktoren in ihren systematischen, soziotechnischen Zusammenhängen nachvollziehen und (4) Gestaltungsentwürfe kontinuierlich umsetzen und rasch testen (Norman, 2019).

Diese Gestaltungsprinzipien bildet Donald Norman auf ein Modell für HCD-Gestaltungsprozesse ab, das er in seinem einschlägigen Buch *The Design of Everyday Things* vorstellt (Norman, 1988, 2013) und das sich aus den vier Aktivitäten Observation, Idea Generation, Prototyping und Testing zusammensetzt (siehe Abb. 8).

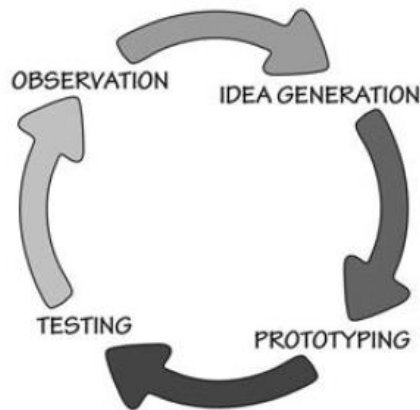


Abbildung 8: Ein menschzentrierter Gestaltungsprozess nach Donald Norman (2013, S. 222)

Die Observation beschreibt dabei die Tätigkeit, Menschen, die ein System nutzen werden, im realen sozialen Kontext zu beobachten, um ihre Ziele und Bedürfnisse zu erheben. Basierend auf dem daraus resultierenden umfassenden Verständnis für ihre Anforderungen kann die Problemdefinition für ein IT-System stattfinden. In der Idea Generation, kurz Ideation, werden erste Ideen und Lösungsansätze entworfen, um die ermittelten Nutzungsanforderungen zu adressieren. Dies ist ein kreativer Prozess, bei dem möglichst viele Lösungen und Ideen entwickelt werden sollen, ohne zunächst mögliche Limitierungen zu beachten oder Ideen vorschnell abzulehnen oder zu verwerfen. Im nächsten Schritt, dem Prototyping, werden sämtliche entwickelten Ideen in Form von wenig aufwändigen Prototypen realisiert. Diese ermöglichen eine erste konkretere Überprüfung ihres Gehalts zur Problemlösung sowie ihres Potenzials zur Aufgabenangemessenheit. Sie haben zudem eine gewisse Kontrollfunktion für den Gestaltungsprozess, denn sie repräsentieren das Verständnis der Nutzungsanforderungen durch die Gestalter und ermöglichen so eine Schärfung bzw. Nachjustierung des Anforderungsverständnisses. Die Prototypen werden anschließend beim Testing durch eine Auswahl repräsentativer Stakeholder getestet, um vertiefende Informationen über das tatsächliche Nutzungsverhalten und die Angemessenheit des Systems zu gewinnen. Idealerweise sind die Gestalter eines Systems auch in diese Phase des Gestaltungsprozesses in beobachtender Funktion involviert. Dadurch markiert das Testing bereits den Beginn des nächsten Gestaltungszyklus. Alle vier Aktivitäten werden im Gestaltungsprozess zur kontinuierlichen Verfeinerung und Verbesserung des Systems iteriert. Dieses Verständnis von MZG prägen Donald Normans Arbeit in den 2000er- und 2010er-Jahren durchgängig (M. W. Meyer & Norman, 2020; Müller et al., 2017; Norman, 2002, 2017).

Auch im Jahr 2024 ist Donald Norman noch zu Fragestellungen der MZG aktiv. Seine Positionen zur MZG haben sich kontinuierlich weiterentwickelt, denn auch er hat, ähnlich wie Joseph Weizenbaum, erkannt, dass die menschliche Natur und besonders das Zusammenwirken von Menschen in Gesellschaftsgefügen Komplexitäten besitzt, die er in seiner bisherigen Arbeit zur MZG vernachlässigt hat. Diesen größeren Zusammenhängen widmet er sich gegenwärtig in seiner erweiterten MZG-Interpretation, der er die Bezeichnung *Humanity-centered Design* (Norman, 2022, 2023; Norman & Euchner, 2023) verleiht.

### 2.2.4 Fokussierung von Nutzerpraktiken und Nutzungskontexten (Lucy Suchman)

Die Technikanthropologin Lucy Suchman trägt mit ihrer Arbeit dazu bei, dass der Nutzungskontext als ein wichtiger Gestaltungsfaktor für die IT-Gestaltung anerkannt wird. Als KI- und HCI-Forscherin am Xerox Palo Alto Research Center und darauffolgend als Professorin für Anthropology of Science and Technology der Lancaster University widmet sich ihre Arbeit der Untersuchung von sozialen Praktiken und angewandtem Wissen, die bei der Nutzung von IT-Systemen wirksam werden. Ähnlich wie bei Donald Norman ist der Ausgangspunkt für Lucy Suchmans Untersuchungen die Fragestellung, weshalb es bei der Interaktion mit IT-Systemen immer wieder zu Missverständnissen und Schwierigkeiten kommt. Auch sie stellt fest, dass IT-Systeme oftmals zu wenig selbsterklärend und intuitiv bedienbar sind. Dies führt sie zurück auf die Diskrepanz zwischen (1) den Annahmen zur Durchführung von Aufgaben, die in die Gestaltung von IT-Systemen einfließen, (2) den tatsächlichen Intentionen von Nutzern und (3) den kontextspezifischen Bedingungen, die Nutzer in der Praxis vorfinden. Sie beschreibt, dass Menschen für die Durchführung von Aufgaben mit IT-Systemen Pläne schmieden, diese Pläne in konkreten Interaktionssituationen jedoch nicht strikt befolgen, sondern entsprechend den vorgefundenen Gegebenheiten adaptieren. Diesem Phänomen widmet sie den Titel ihrer einschlägigen Arbeit *Plans and situated actions* (Suchman, 1985, 1999), in der ihr als Beispiel ein Xerox-Kopiergerät und dessen UI dient. Lucy Suchman leitet aus ihren Beobachtungen ab, dass Menschen in der Interaktion mit IT-Systemen auf ihre Pläne mehr als Ressource wie als streng zu befolgende Instruktionen zurückgreifen. Ferner argumentiert sie, dass Nutzer ungeachtet ihres technischen Wissenstandes im Aufeinandertreffen mit einem technischen Artefakt sich zusätzlich mit der Aufgabe eines „making sense of a new artifact [as] an inherently problematic activity“ (Suchman, 2007, S. 9) konfrontiert sehen. Ihre Darstellung von Interaktionsprozessen als situierte Aneignungs- und Interpretationshandlungen hat tiefgreifende Implikationen für die Debatten zur IT-Gestaltung Anfang der 2000er-Jahre, denn Lucy Suchman kritisiert damit die in den Kognitionswissenschaften wie auch in der KI-Forschung seit Ende der 1980er-Jahre vertretenen Annahmen zur menschlichen Wissensrepräsentation (Adelson, 2003). Ihre Erkenntnisse zeigen ferner, dass das in der IT- und KI-Gestaltung übliche Vorgehen des Modellierens menschlicher Handlungspläne und die Verwendung dieser Modelle als primärer Bezugspunkt für die interne Logik sowie das Interaktionskonzept von IT-Systemen nicht länger haltbar sind. Dies stützt sie auf die von ihr festgestellten Asymmetrien zwischen den formalen Kommunikationspfaden techni-

scher Systeme und den Dynamiken im (zwischen-) menschlichen Kommunikationsverhalten. Für Lucy Suchman ist ein wichtiger Erfolgsfaktor zwischenmenschlicher Interaktion die Erkennung und Klärung sich unterscheidender Verständnisse von Sachverhalten, die eine „kind of presence to the unfolding situation of interaction, not available to the machine“ (Suchman, 2007, S. 12) beinhalten. Wie Donald Norman spricht sie sich deshalb auch für eine stärkere Fokussierung der Gestaltung von UIs im Sinne einer „wesentlichen Voraussetzung gelingender Interaktion“ (Pentenrieder & Weber, 2020, S. 217) aus, auch wenn sie die Position vertritt, dass kein noch so gut gestaltetes Interface die Tätigkeit der Aneignung vollständig eliminieren wird (ACM SIGCHI, 2016).

Die Arbeit Lucy Suchmans hat neben solchen erkenntnistheoretischen Beiträgen auch Auswirkungen auf die methodischen Praktiken der IT-Gestaltung, die für die Entwicklung des MZG-Ansatzes von entscheidender Bedeutung sind. Denn zur Durchführung ihrer Forschung wendet Suchman empirische Methoden der Ethnomethodologie (Garfinkel, 1967) sowie der Konversationsanalyse (Sacks et al., 1978) an und verbindet damit erstmalig im Feld der IT-Gestaltung ethnografische Studien mit Fragestellungen der Nutzung von IT-Systemen (siehe Abb. 9).

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale
S1	[Points] Right there.		
S2	[Hands on latch]		
S1	"And lift up to the left." [looks to S2, who struggles with the latch]		
	"Lift up and to the left."		
S2	[Still struggling]		
S1	Okay::		
S2	Pu::ll, and lift up to the left. [Looks at picture] Oh, the whole thing		
S1	Yea.		
S2	lift up and to the left.		
	RAISES DOCUMENT HANDLER		
		DISPLAY 3	
S1	"Place your original face down, [Passes journal to S2] on the glass, centered over the registration guide." [looks to machine] Got that? [pause] Want me to read it again?		Instructions for placing document and closing document cover.
S2	Um:: I'm just trying to figure out what a registration guide is, but I guess that's this, um:		
S1	[Looking over her shoulder] Yea:		
S2	centered over this line thingy here.		

Abbildung 9: Von Lucy Suchman erstelltes Template zur kombinierten Konversations- und Systemzustandsanalyse (Suchman, 1985, S. 80)

Die Arbeit von Lucy Suchman zeigt damit, dass es ethnografische Untersuchungen von Nutzungspraktiken in realen Nutzungskontexten ermöglichen, den Prozess von Bedeutungszuschreibungen von Menschen zu IT-Systemen nachzuvollziehen. Ihr Fokus auf konkrete Mikropraktiken und -prozesse und ihre empirische Arbeitsweise haben Lucy Suchman zu einem zentralen Referenzpunkt in der ethnografischen Technikforschung und -gestaltung werden lassen. So gilt sie als Wegbereiterin der Workplace Studies, die die Interaktion mit als intelligent bezeichneten Artefakten am Arbeitsplatz untersuchen (Bergmann, 2004).

### 2.3 Gestaltungsprinzipien

Wie aus der Beleuchtung der MZG-Historie ersichtlich wird, sind die Bemühungen zur Verknüpfung theoretischer Erkenntnisse und methodischer Vorgehensweisen ein bedeutender Bestandteil der MZG-Tradition. Die Ergebnisse dieser Bemühungen stellen sich auch in Form von Gestaltungsprinzipien dar, die Gestaltungsverantwortlichen Orientierung und Anleitung zur Planung, Umsetzung und Evaluation menschzentrierter Gestaltungsprozesse bieten sollen. Die MZG-Literatur verzeichnet mittlerweile zahlreiche Vorschläge zu Gestaltungsprinzipien (z.B. Notess & Blevis, 2004; Oviatt, 2006; Zarte et al., 2020), die aufgrund ihrer Anzahl eine Orientierung für potenziell interessierte Gestaltungspraktiker erschweren. Für sie ist es daher hilfreich, wenn eine zentrale Referenz zur Hand ist, die die wesentlichen MZG-Zusammenhänge darlegt und auch auf Gestaltungsprinzipien eingeht. Die Funktion eines solchen Referenzdokuments erfüllt gegenwärtig die Norm DIN EN ISO 9241-210 (DIN EN ISO 9241-210:2020-03, 2020), in der die grundsätzlichen Motivationen der MZG mit Hinweisen und Instruktionen zur Realisierung menschzentrierter Gestaltungsprozesse verständlich dargelegt werden. Die Norm hat sich international als belastbare Referenz für die Gestaltungspraxis etabliert (Bevan et al., 2015; Jokela et al., 2003; Schneider, 2008), was auch darauf zurückzuführen ist, dass in ihr die Zusammenhänge zu Tätigkeiten des Projektmanagements sowie die wirtschaftlichen Vorteile der MZG ersichtlich werden. Die wesentlichen Inhalte der Gestaltungsprinzipien beinhalten (1) die Erfassung des Nutzungskontexts, (2) die Beteiligung von Nutzern, (3) iterative Gestaltungsprozesse, sowie (4) Gestaltungsteams, die über interdisziplinäre Expertise verfügen (siehe Abb. 10).





Abbildung 10: Gestaltungsprinzipien der MZG, abgeleitet aus DIN EN ISO 9241-210:2020-03

### 2.3.1 Erfassung des Nutzungskontexts

Ein zentrales Gestaltungsprinzip der MZG ist die Erarbeitung eines umfassenden Verständnisses über den Nutzungskontext, der die Zusammenhänge zwischen den Anforderungen und Zielen direkter und indirekter Stakeholder beschreibt. Auch beschreibt der Nutzungskontext die Bedingungen unter denen ein System zum Einsatz kommen wird (DIN EN ISO 9241-210:2020-03, 2020, S. 14; Hernández-Ramírez, 2020; Hyysalo & Johnson, 2015; Ritter et al., 2014; Roschuni et al., 2013). Ziel ist das Verstehen der Aufgaben, die Menschen mit einem System durchführen werden, sowie die Freilegung relevanter, tieferliegender Motivationen für die Interaktion:

*„We aim to understand their particular activities, beliefs, preferences, emotions, motivations, troubles, and environments. We want to see users’ messy realities; to understand their mistakes, short-cuts, and work-arounds; and to learn about the personal significance and meaning they attach to their activities and product interactions“ (Van Pelt & Hey, 2011, S. 690).*

Die Erfassung des Nutzungskontexts umfasst neben der Beschreibung soziotechnischer Rahmenbedingungen und Beziehungen auch die Betrachtung ökonomischer, organisationaler, physischer und kultureller Aspekte, die sich auf die Nutzung eines IT-Systems auswirken (können). Je nach Gestaltungskontext gilt es die relevanten Aspekte zu identifizieren, was die Ermittlung von Räumlichkeiten, Arbeitspraktiken, Fachterminologien oder auch Organisationskulturen beinhalten kann. Zur Erfassung solcher Gestaltungsfaktoren kann es hilfreich sein, bestehendes Wissen zum Nutzungskontext aufzugreifen, welches beispielsweise in Form von Leitbildern, Prozessbeschreibungen oder Architekturplänen vorhanden ist. Auch kann es instruktiv sein, Berichte und Ergebnisse zuvor durchgeführter IT-Gestaltungsprojekte zu sichten, um die Rahmenbedingungen sowie bestehende Soft- und Hardwarestrukturen eines Nutzungskontexts nachzuvollziehen. Zusätzlich zur Ermittlung intendierter Einsatzszenarien kann es von großem Wert sein, in Erfahrung zu bringen, was Stakeholder mit der Nutzung eines Systems verhindern möchten oder welche Bedenken und Vorbehalte sie zu einer IT-gestützten Lösung eines Problems hegen.

Ein in diesem Sinne angemessenes und möglichst vollständiges Bild vom realen Nutzungskontext stellt eine zentrale Informationsquelle für den menschenzentrierten Gestaltungspro-



zess dar. Es bildet die Grundlage und den permanenten Bezugspunkt für den gesamten Gestaltungsprozess, an dem sämtliche Gestaltungsaktivitäten ausgerichtet werden. Vom Nutzungskontext werden die Anforderungen an ein System abgeleitet, weshalb sich eine angemessen andauernde, aktive Auseinandersetzung mit Stakeholdern, ihren Praktiken und Umgebungen empfiehlt. Um diesem Gestaltungsprinzip Rechnung zu tragen stellt deshalb auch in menschenzentrierten Gestaltungsprozessen das Requirements Engineering (Macaulay, 2012; Pohl, 1996) als Prozess zur Definition und Dokumentation von Anforderungen eine zentrale Maßnahme dar. Typische Herausforderungen der Erfassung des Nutzungskontexts beinhalten den Umgang mit stark divergierenden Anforderungen heterogener Stakeholdergruppen sowie die Definition der Systemgrenzen. Ferner müssen Umgangswege zur Dynamik von Einsatzräumlichkeiten gefunden werden und auch die vorgreifenden oder rückwirkenden Auswirkungen eines Systems auf benachbarte Prozesse stellen häufig Herausforderungen dar.

### 2.3.2 Nutzerbeteiligung

Ein weiteres zentrales Gestaltungsprinzip der MZG ist die Einbindung und Beteiligung von Stakeholdern in den Gestaltungsprozess. Dies erfolgt in der Regel durch eine aktive Mitwirkung von Nutzern in der Gestaltung, z.B. durch die Evaluation von Prototypen oder Systemversionen und variiert in Form und Häufigkeit je nach konkretem Projekt. Die Menschen, die später mit einem System arbeiten werden, verfügen über fachspezifisches Domänen- und Erfahrungswissen und kennen die Charakteristika des Nutzungskontexts, wie beispielsweise Hierarchien, Abläufe und Terminologien, weshalb sie eine für den Prozess wertvolle Informationsquelle darstellen. Durch die Konsultation und den Austausch mit Nutzern kann solch intrinsisches Wissen freigelegt werden, um dessen Bedeutung für den Gestaltungsprozess zu ergründen und um relevante Aspekte für die Gestaltung aufzugreifen und während der Gestaltung zu berücksichtigen. Gestalter und Nutzer kommen miteinander in Kontakt und erhalten wertvolle, für beide Seiten dienliche Einblicke in die Realitäten „beider Welten“. Dadurch kann auch das gegenseitige Vertrauen sowie das Vertrauen in das spätere System gestärkt werden, denn Nutzer können erfahren, welche direkten Auswirkungen ihr Feedback auf die Gestaltung eines Systems hat.

In der Evaluation eines Systems durch Nutzer liegt das zentrale Anliegen dieses Gestaltungsprinzips, denn sie vermindert das Risiko einer fehlgeleiteten Systemgestaltung. Zentral für die MZG ist die Erhebung und spätere Orientierung an den expliziten und impliziten Anforderungen der Stakeholder, wozu die Evaluation eines Systems zu bestimmten Zeitpunkten im Prozess eine zusätzliche absichernde Funktion einnimmt. Sie ermöglicht qualifiziertes Feedback unter realitätsnahen Nutzungsbedingungen einzuholen, das Auskunft zur Aufgabenangemessenheit, Usability und weiteren relevanten Aspekten hervorbringt. Sie zeigt, inwiefern ein System die Anforderungen erfüllt und inwiefern es die Fähigkeiten und Ziele von Nutzern berücksichtigt. Da die Evaluation zu einem Zeitpunkt stattfindet, wo ein System noch formbar und veränderbar ist, können möglicherweise hervortretende Unsicherheiten, Anforderungsveränderungen oder andere erkennbar relevante Dynamiken im weiteren Gestaltungsverlauf adressiert werden.

### 2.3.3 Iterative Gestaltungsprozesse

Eng verknüpft mit der partizipativen Einbindung von Nutzern ist das Gestaltungsprinzip der iterativen Gestaltungsprozesse. Iterationen stellen Prozessschleifen oder -zyklen dar, in denen Gestaltungsaktivitäten (mehrfach) wiederholt werden, um sich dem angestrebten Gestaltungsergebnis sukzessive anzunähern. Hintergrund dieses Prinzips ist der Erfahrungswert, dass erste Gestaltungslösungen häufig nicht ausreichend angemessene oder zufriedenstellende Resultate darstellen. Dies kann verschiedene Ursachen haben, beispielsweise indem Charakteristika des Nutzungskontexts die Umsetzung einschränken, oder eine bislang weniger im Fokus stehende Nutzergruppe und ihre Anforderungen sich als hoch relevant für den Gestaltungsprozess herausstellen. Durch die Arbeit mit Iterationsschleifen wird versucht, solchen Dynamiken in der Realität Rechnung zu tragen und vorzubeugen, möglichst produktiv mit ihnen zu arbeiten und in Summe dem Gestaltungsprozess eine höhere Flexibilität zu verleihen. Das Prinzip iterativer Prozesse setzt eine Akzeptanz für den Umstand voraus, dass im Laufe des Gestaltungsprozesses wiederholt Änderungen am System durchgeführt werden müssen. Auch erfordert es, die Kommunikation der voraussichtlich häufigen Systemveränderungen möglichst zu Beginn eines Gestaltungsprojekts allen Beteiligten transparent zu kommunizieren. Dies bedeutet häufig ein aufwändiges Projektmanagement und fordert von allen an der Gestaltung Beteiligten ein hohes Maß an Ausdauer. Erkenntnisse einer Iterationsschleife können beispielsweise zeigen, dass bestimmte Features eines Systems in deutlich reduzierterem als bislang antizipierten Umfang Bestandteil des Systems sein werden oder das kritische Systemkomponenten oder Funktionen noch deutlich stärker ausgebaut werden müssen. Die Korrektur von Annahmen sowie ein pragmatischer Umgang mit Mehrdeutigkeiten zeichnen dadurch den gesamten Gestaltungsprozess aus.

### 2.3.4 Interdisziplinäre Gestaltungsteams

Da sich die MZG mit den soziotechnischen und weiteren relevanten Aspekten für die Gestaltung von IT-Systemen auseinandersetzt, sieht ein weiteres ihrer Gestaltungsprinzipien das Vorhandensein interdisziplinärer Expertise in den verantwortlichen Gestaltungsteams vor. Dies dient der angemessenen Begegnung der verschiedenen fachübergreifenden Faktoren sowie der konstruktiven Unterstützung der Findung von Kompromissentscheidungen (Pagel et al., 2022). Je nach Nutzungskontext können dafür zusätzliche Expertisen aus Fachbereichen wie der Produktgestaltung, Systemanalyse, Marketing, Vertrieb, Support und weiteren hilfreich sein. Durch die Beteiligung von Nutzern zählen auch sie zum interdisziplinären Charakter, den ein MZG-Prozess auszeichnet. Vorteil der Interdisziplinarität ist neben der erweiterten Kompetenzbasis eine höhere Problemlöseaktivität im Team, sowie die Erhöhung der Sensibilität für Möglichkeiten und Grenzen der vertretenen Kompetenzbereiche. Herausforderungen können dadurch entstehen, wenn zu systemkritischen Aspekten unterschiedliche Ansichten vorherrschen, die nicht auf gleiche Weise in Gestaltungsentscheidungen berücksichtigt werden können.

## 2.4 Zusammenfassung

Die Ideen, die im Ansatz der MZG verankert sind, sind das Ergebnis jahrzehntelanger Bemühungen, die durch eine intensive Zusammenarbeit über disziplinäre Grenzen hinweg geprägt wurden. Kennzeichen der MZG sind die Umkehrung der Priorisierung von technischen hin zu menschlichen Anforderungen in Gestaltungsprozessen, sowie die Aufhebung der gedanklichen und praktischen Trennung von Gestaltungs- und Nutzungstätigkeiten. Besonders an der MZG ist auch ihre inhärente Verknüpfung von Gestaltungspraxis und Gestaltungstheorie, in der Gestaltungsentscheidungen basierend auf Nutzerwissen und Nutzerfeedback getroffen werden. Diese Vorgehensweisen besitzen auch im Jahr 2024 Relevanz und werden in verschiedenen Communities praktiziert und weiterentwickelt. Die Ideen der Wegbereiterinnen und Pioniere konnten sich in der Praxis durch die Standardisierung von Gestaltungsprinzipien bewähren, besitzen jedoch nach wie vor Potenziale zur Weiterentwicklung. Denn auch wenn die Ideale der MZG für die Gestaltung diverser Technologien Relevanz besitzen, wurden sie primär entwickelt, um auf die Gestaltung von Softwaresystemen anwendbar zu sein. Dadurch sind der MZG Grenzen gesetzt, die angesichts neuer Entwicklungen im gesellschaftlichen Miteinander und der technischen Möglichkeiten evident werden.

## 3 Menschzentrierte KI-Gestaltung

Die Ideen und Vorschläge der MZG erfahren erneut erhöhte Aufmerksamkeit in Kontexten zur Gestaltung von KI-Systemen. Anlass zur Steigerung des Interesses an der MZG sind die Fortschritte der KI-Forschung und -Entwicklung, die alte und neue Fragen zu sozialen und ethischen Implikationen von KI-Systemen in den Vordergrund öffentlicher Diskurse (z.B. Clark & Chamorro-Premuzic, 2024; Datenethikkommission der Bundesregierung, 2019; European Commission, 2019; Magdans, 2024) und fachlicher Diskurse (z.B. Bélisle-Pipon et al., 2023; Hagendorff, 2020; Heinlein & Huchler, 2024; Jobin et al., 2019; Tiwari, 2023; Whittlestone et al., 2019) rücken. Diese befassen sich mit den verschiedenen Herausforderungen und Fragen, die die mit der menschenzentrierten Gestaltung und Nutzung der KI-Technologie verbundenen Zielsetzungen aufwerfen. Die grundlegende Prämisse dieser Diskurse wird in dem in 2024 verabschiedeten EU AI Act treffend zusammengefasst:

*„AI should be a human-centric technology. It should serve as a tool for people, with the ultimate aim of increasing human well-being“* (European Parliament, 2024, S. 6).

Mit der Realisierung dieser Aufgabe befassen sich verschiedene Communities, zu welchen auch die vergleichsweise junge interdisziplinäre Community der HCAI zählt, die Ende der 2010er-Jahre aus der HCI-Community hervorgegangen ist. Ihre Mitglieder untersuchen als eine zentrale Fragestellung, inwieweit die Theorie und Methodik der MZG für eine MZKG adaptiert werden kann. Wie sich die Forschungslandschaft des Feldes zusammensetzt (siehe Kap. 3.2) und welche Modelle zur Weiterentwicklung der MZG diskutiert werden (siehe Kap. 3.3), gebe ich nachfolgend in Form eines Überblicks wieder. Abschließend stelle ich kurz einige für das Feld HCAI charakteristische Debatten dar, die das Feld in jüngerer Vergangenheit geprägt haben und perspektivische Relevanz besitzen (siehe Kap. 3.4). Auf diese Weise wird es mir möglich, die in HCAI vorherrschenden Interpretationen der MZKG darzustellen, auf die ich in den nachfolgenden Kapiteln Bezug nehmen werde.

### 3.1 Einführung

KI-Systeme haben viele Bereiche des gesellschaftlichen Lebens geprägt. Moderne KI-Systeme werden inzwischen dazu eingesetzt, Krankheitsbilder zu identifizieren (Bhasker et al., 2023), neue Formen der Pflanzenzüchtung zu erschließen (Harfouche et al., 2019), oder auch um bei der Organisation des beruflichen und privaten Alltags zu unterstützen (Kepuska & Bohouta, 2018). Weitere Anwendungen finden sich in Bereichen wie dem Finanzwesen (Cao, 2023), dem Mobilitätssektor (Namatherdhala et al., 2022) oder auch der Fertigungsindustrie (Wennker, 2020), wo der Einsatz der Systeme neben ökologischen und ökonomischen besonders auch mit sozialen Nutzeffekten verbunden wird. Inspiration und Motivation vieler Anstrengungen der KI-Gestaltung ist die Versatilität des menschlichen Gehirns (Krauss, 2023; Teich, 2020), dessen Fähigkeiten zur Analyse von Daten und Verknüpfung von Informationen sie technisch nachzubilden und zu erweitern sucht. Dazu

verwenden KI-Gestalter daten- und rechenintensive Verfahren wie Machine Learning, Pattern Recognition oder Neuronale Netze, die die KI-Systeme zur Durchführung von für menschliche Verhältnisse komplexe, eintönige oder zu gefährliche Aufgaben befähigen sollen (Peeters et al., 2021). Wie in anderen Bereichen der TG bedeutet auch die Gestaltung von KI-Systemen unweigerlich eine Gestaltung von Mensch-Technik-Beziehungen (Herrmann & Pfeiffer, 2023; Kluge et al., 2021; Latniak et al., 2023). Die in KIGP getroffenen Entscheidungen definieren neben technischen Aspekten maßgeblich, wie die Einbettung von KI-Systemen in soziale Kontexte erfolgen kann. Dies betrifft im Besonderen die Aneignung der Systeme und die aus ihrer Nutzung resultierenden Adaptionen und Neubildungen sozialer Praktiken. Angesichts der zunehmend steigenden Anzahl von Anwendungsbereichen für KI-Systeme wird erkennbar, dass KI-Systeme sich auch zukünftig auf die Leben vieler Menschen auswirken werden. Angesichts solcher Tendenzen steigt auch die mit der Gestaltung von KI-Systemen verbundene soziale Verantwortung, die eine reflektierte Auseinandersetzung mit den Aktivitäten des Gestaltungshandelns erfordert und eine besonnene Auswahl der Methoden und Instrumente nahelegt, die zur Herbeiführung von Gestaltungsentscheidungen verwendet werden. Für Gestalter und Entscheidungsträger in KIGP hat die Auseinandersetzung mit den sozialen und ethischen Implikationen ihrer Arbeit entsprechend an Bedeutung gewonnen (Dwivedi et al., 2021; Floridi et al., 2021; Fosso Wamba et al., 2021; Makridakis, 2017; Wittpahl, 2019).

Begleitet und befeuert wird der akademische Diskurs kontinuierlich durch lebensweltliche Geschehnisse, wie etwa der Veröffentlichung der Large Language Model-Anwendung ChatGPT3 (Abdullah et al., 2022; J. G. Meyer et al., 2023) oder dem Text-to-Image Generator Midjourney (Chiu, 2023; Wasielewski, 2023) im Jahr 2022. Beide KI-Systeme verzeichneten bereits nach kurzer Zeit ihrer Veröffentlichung mehrere Millionen Nutzer (Ploennigs & Berger, 2023; Trust et al., 2023), veränderten Branchendynamiken und Jobprofile in Bereichen wie der Softwareentwicklung, der Architektur, des Grafikdesigns oder des Bildungswesens (Chiou et al., 2023; Firth & Triche, 2023; Jiang et al., 2023; S. Lee, 2023; Steinfeld, 2023) und rückten rasch in den Vordergrund der medialen Aufmerksamkeit und öffentlichen Debatte (siehe Abb. 11).



Abbildung 11: Werbeplakat mit der Aufschrift „Hey Chat GPT, finish this building...“ (Werbeagentur IMPACT, 2023)

Solche Beispiele unterstreichen, weshalb das Anbringen einer besonderen Sorgfalt und Rücksicht für die Eigenschaften und Werte menschlichen Zusammenlebens zu einer immer bedeutsameren Aufgabe für die KI-Gestaltung werden (V. Dignum, 2017; Floridi, 2021; Jobin et al., 2019; Umbrello & van de Poel, 2021; van de Poel, 2020; Waefler & Schmid, 2020). Wie die auf gesellschaftlicher und politischer Ebene geäußerten Forderungen nach menschenzentrierten KI-Systemen erfüllt werden können, wird in verschiedenen Fachbereichen erörtert. Hinsichtlich der konkreten gestalterischen Herausforderungen hat sich das Feld HCAI zu einem zentralen Austragungsort etabliert (Ozmen Garibay et al., 2023; Riedl, 2019; Shneiderman, 2020, 2022). In HCAI liegen die Schwerpunkte auf der multidisziplinären Auseinandersetzung mit den theoretischen Aspekten der MZKG sowie auf der Operationalisierung erarbeiteter theoretischer Erkenntnisse für die Gestaltungspraxis.

## 3.2 Überblick der Forschungslandschaft

Waren in den 1980er-Jahren die in der IT-Gestaltung beobachtbaren technikzentrierten Tendenzen Hauptmotivation für die Entwicklung von Ansätzen wie der MZG, so weist auch die Motivation zur Gründung des Feldes HCAI vergleichbare Merkmale auf. So lässt sich auch beim Einsatz von KI-Systemen feststellen, dass Fehler auftreten (Yampolskiy, 2019), dass bei ihrer Gestaltung der Nutzungskontext ignoriert wird (Fischer, 1995) oder dass sie für menschliche Kollaborationsweisen zu unflexibel sind (Fan & Yen, 2007). Einen guten Eindruck über die Defizite und Risiken moderner KI-Systeme vermittelt beispielsweise die AI Incident Database (Responsible AI Collaborative, 2024). Sie verzeichnet Mitte 2024 knapp 700 Einträge zu Vorfällen mit (Beinahe-)Schadensfällen, die durch den Einsatz von KI-Systemen verursacht wurden. Hier finden sich Berichte zu durch Fehlfunktionen autonomer Fahrzeuge und Chirurgie-Roboter verursachte Verletzungen und Todesfälle, oder die Verbreitung von Falschinformationen in Medienberichten und Bilddatenbanken durch von KI-Systemen generiertes Text- und Bildmaterial. Die Positionen und Argumente, die die Entwicklung der MZG motivierten, besitzen somit unzweifelhaft auch Gültigkeit für die KI-Gestaltung und wiegen für sie erheblich schwerer (Xu, Dainoff, et al., 2023), da KI-Systeme auf globaler Ebene in unterschiedlichsten Lebensbereichen eingesetzt werden.

### 3.2.1 Einrichtungen

Um solchen und weiteren Herausforderungen zu begegnen, haben sich zahlreiche Bemühungen im Feld HCAI konzentriert. Einen guten Eindruck über die in HCAI stattfindenden Aktivitäten vermittelt der Blick auf die organisationale Ebene des Forschungsfelds, wo beispielsweise die Stanford University einen Lehrstuhl mit dem Namen Human-Centered Artificial Intelligence ins Leben gerufen hat (Stanford Institute for Human-Centered Artificial Intelligence, 2023). Neben Stanford haben auch zahlreiche weitere Universitäten vergleichbare Lehrstühle und Forschungsgruppen zu HCAI eingerichtet, die sich weltweit mit verschiedenen inhaltlichen Schwerpunkten auseinandersetzen (siehe Tab. 1).

Lehrstuhl oder Forschungsgruppe	Bildungsreinrichtung
Berkman Klein Center for Internet and Society	Harvard University
Center for Human, Artificial Intelligence, and Robot Teaming	Arizona State University
Center for Human-Compatible AI	University of California
Center for Human-Computer Interaction + Design	Northwestern University (Illinois)
Center for Responsible AI	New York University
Center for Responsible Machine Learning	University of California
Center for Socially Responsible Artificial Intelligence	Penn State University
Centre for AI Technology for Humankind	National University of Singapore
Centre for Machine Intelligence	University of Sheffield
Chicago Human+AI Lab	University of Chicago
Ethics of AI Lab	University of Toronto
Human Centered AI	Monash University
Human Centred Technology Research Centre	University of Canberra
Human(e) AI	University of Amsterdam
Human-AI Interaction	University of British Columbia
Human-centered AI	Utrecht University
Human-Centered AI	Universität Kassel
Human-Centered AI Lab	Universität für Bodenkultur Wien
Human-Centered Artificial Intelligence	Otto von Guericke Universität Magdeburg
Human-Centered Artificial Intelligence Lab	University of Seoul
Human-centric Artificial Intelligence Centre	University of Technology Sidney
Humanising Machine Intelligence	Australian National University
Humanity Centered Robotics Initiative	Brown University
Institute for People-Centered Artificial Intelligence	University of Surrey
Institute for Assured Autonomy	Johns Hopkins University
Institute for Ethics in AI	University of Oxford
Institute for Multimedia and Interactive Systems	Universität zu Lübeck
Internet Institute, Future of Humanity Institute	University of Oxford
ISOVIS group	Linnaeus University
iVis group	Linköping University
Lehrstuhl für Menschzentrierte Künstliche Intelligenz	Augsburg University
Leverhulme Centre for the Future of Intelligence	University of Cambridge
Research Cluster of Human-Centered AI	Aalborg University
Tech3Lab	HEC Montreal

Tabelle 1: Auswahl von HCAI-Forschungsgruppen und -Lehrstühlen im Jahr 2024 (Einträge teilweise aus Shneiderman & Mengan, 2024)



Auch bieten verschiedene Bildungseinrichtungen inzwischen Master-Studiengänge im Fach HCAI an, wie beispielsweise das Kooperationsprogramm Human-Centered Artificial Intelligence Master der Universitäten Budapest, Utrecht, Neapel und Dublin (The Human-Centred AI Master's Programme and Community, 2024), das Masterprogramm der Universität Göteborg (University of Gothenburg, 2024) oder das Master-Kooperationsprogramm der Universitäten Milano-Bicocca und Pavia (Università degli Studi di Milano Statale, 2024). Die Bemühungen internationaler Zusammenarbeit im Feld HCAI werden darüber hinaus auch im Forschungsnetzwerk Humane AI Net gefördert (Humane AI Net, 2024). Dieses hat sich dem Ziel verschrieben, einen Beitrag zu einer KI-Gestaltung nach europäischen ethischen, sozialen und kulturellen Normen und Werten zu leisten. Sämtliche führenden europäischen Forschungseinrichtungen sind Teil dieses Netzwerks, was beispielsweise durch die Beteiligung des deutschen Forschungszentrums für Künstliche Intelligenz DFKI, der ETH Zürich oder der Technischen Universität Delft deutlich wird.

### 3.2.2 Themenfelder

All diese und weitere Kooperationen und Einrichtungen tragen dazu bei zu ergründen, was eine MZKG bedeuten kann und wie sie konzeptionell, organisatorisch und technisch realisiert werden kann. Da sich die Community nach wie vor im Aufbau befindet, erweitern sich ihre Grenzen stetig und auch der Themenkatalog des Feldes steigt stetig an. Die Grenzen und Überschneidungen von HCAI zu bestehenden und benachbarten Fachbereichen sind fließend, da diese in HCAI teilweise aufgegriffen und zu ausgewählten Fragestellungen der KI-Technologie neu zusammengeführt werden. Einen guten Eindruck der in HCAI im Vordergrund stehenden Inhalte und ihrer Vielfalt liefern die Ergebnisse einer Umfrage, die auf der Konferenz Hybrid Human-Artificial Intelligence im Jahr 2022 unter den Konferenzteilnehmern durchgeführt wurde. Die Ergebnisse der Umfrage geben wieder, welche Themenfelder die Mitglieder der HCAI-Community für besonders bedeutsam erachten (siehe Abb. 12)

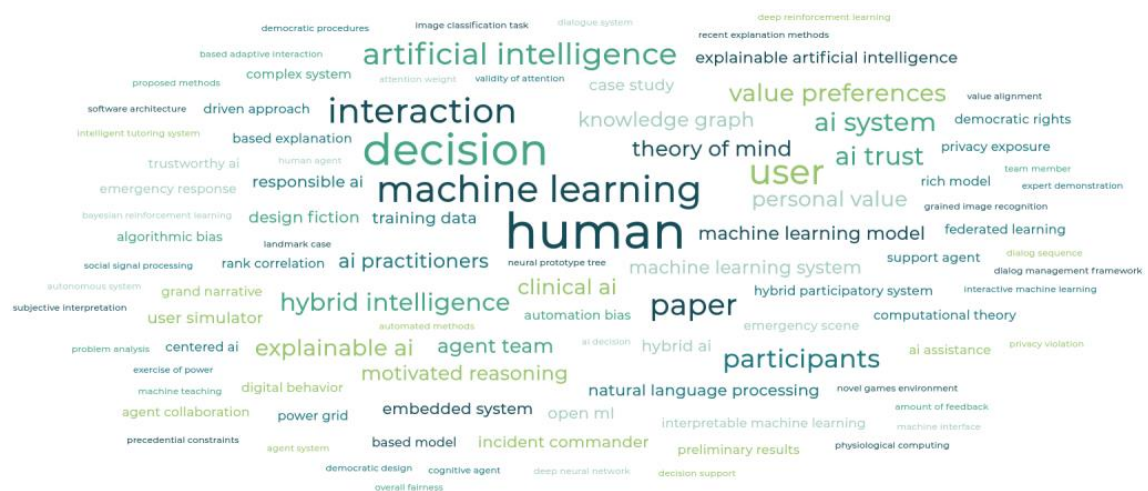


Abbildung 12: Themenübersicht der HCAI-Konferenz in 2022 (Hybrid Human Artificial Intelligence (HCAI), 2023)

Neben technischen Verfahren der KI-Gestaltung sind für die Mitglieder der HCAI-Community Fragen der Ethik von Bedeutung, wie beispielsweise die Einträge zur Priorisierung von Werten in Gestaltungsprozessen (value preferences) oder der Wahrung von Privatsphäre (privacy exposure, privacy violation) zeigen. Auch stehen Fragen zu Methoden der MZKG im Vordergrund, was beispielsweise an der Nennung der Methode des Design Fiction und dem Eintrag „proposed methods“ ersichtlich wird. Auch gehen aus der Umfrage die inhaltlichen Anknüpfungspunkte zu Feldern wie dem Explainable AI, dem Trustworthy AI oder dem Physiological Computing hervor, die ebenfalls noch recht junge Geschichten besitzen. Auf Detailebene stehen dabei in HCAI viele Fragen zu dem Gesichtspunkt im Raum, inwieweit eine MZKG auch dafür Verantwortung tragen kann und muss, ethische, psychologische oder arbeitswissenschaftliche Gütekriterien zu berücksichtigen. Die beiden Autorinnen Capel und Brereton beschreiben den gegenwärtigen Zustand der Debatte treffend mit den Worten:

*„there remains considerable ambiguity about what it means to frame, design and evaluate HCAI“* (Capel & Brereton, 2023, S. 1).

Es handelt sich dabei um einen Findungsprozess von Verantwortungs- und Gültigkeitsgrenzen, der sich auch in 2024 noch in vollem Gange befindet. Die Herausforderungen dieses multidisziplinären Aushandlungsprozesses werden beim weiteren Blick in die Arbeit von Capel und Brereton ersichtlich, die die Forschungslandschaft von HCAI anhand von Publikationsaufkommen kartiert haben (siehe Abb. 13).

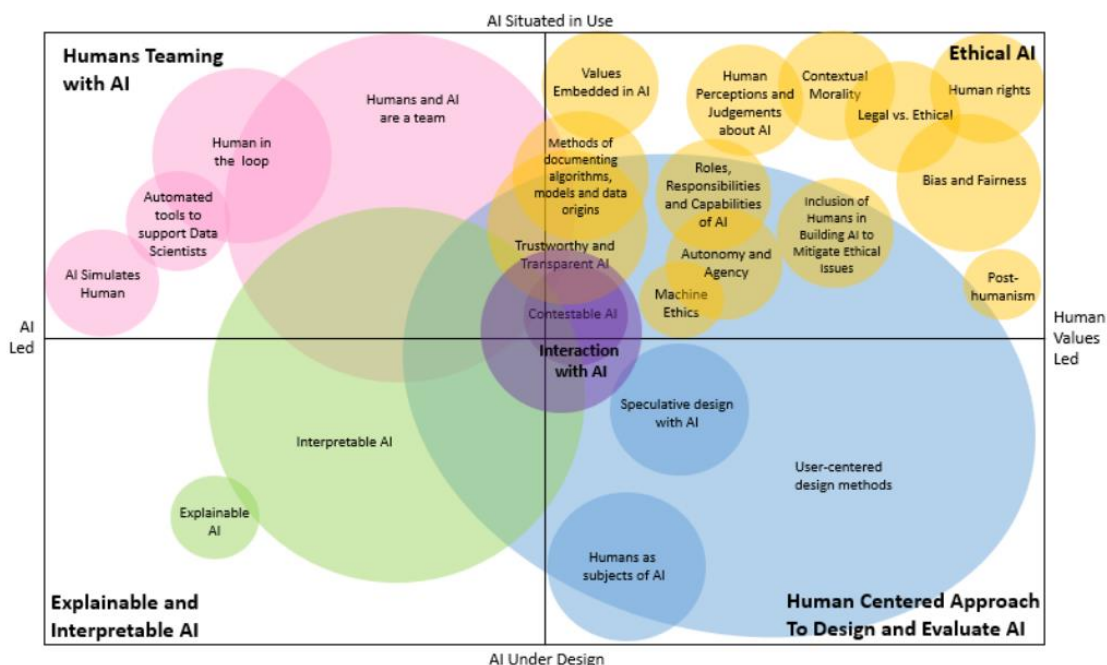


Abbildung 13: Themenfelder in HCAI (Capel & Brereton, 2023, S. 5)

Die Autorinnen teilen das Feld in vier Themenfelder ein, dem Human-AI-Teaming, den bereits erwähnten Feldern der Ethical AI und Explainable AI, sowie der Perspektiven auf die KI-Gestaltung aus Sicht der MZG-Tradition. Hieraus wird erkennbar, wie sehr sich die Themenlandschaft in HCAI im Vergleich zur HCI-Tradition erweitert hat – Fragen zur Ethik und zur Nachvollziehbarkeit sind charakteristisch für Debatten der KI-Gestaltung und erfordern in HCAI das Einnehmen höherer Abstraktionsebenen und die Auseinandersetzung mit Fragestellungen, die auf der reinen Betrachtung der Interaktionsebene von Mensch und Technik (zentrale Expertise in HCI, auch visuell in Abb. 13) nicht zielführend genug behandelt werden können. Besonders relevant für die Forschungsfrage der Dissertation ist dabei die Feststellung der beiden Autorinnen, dass der Mensch in HCAI-Arbeiten stark heterogen konzipiert wird. Hier findet ein Spektrum an Konzepten Anwendung, welche den Menschen z.B. als Subjekt eines KI-Algorithmus betrachten, die aber auch das traditionelle User-Verständnis anwenden sowie verschiedene Konzepte des Menschen als gestalterisch schaffendem Verantwortungsträger. Damit sprechen die Autorinnen eine weitere zentrale Fragestellung für die Community an – Welche Vorstellungen von menschlichem Leben und Handeln werden von den Communitymitgliedern zugrunde gelegt, wenn sie von menschenzentrierter KI-Gestaltung sprechen? Mit anderen Worten: Auf welche Vorstellungen des Menschen beziehen sich die Debatten zur MZKG? Dies stellt eine gewichtige Frage dar, da die möglichen Antworten auf diese Frage nicht zuletzt normative Implikationen aufweisen. Da sich die Frage menschlichem Leben und Handeln widmet, muss bei der Erarbeitung von Antworten zudem eine angemessene Form von Behutsamkeit und Sorgfalt angewendet werden, wie sie etwa in den Diskursen der Technikanthropologie Tradition ist.

### 3.3 Erweiterte Modelle menschenzentrierter Gestaltung

Parallel zu dieser Auseinandersetzung finden in HCAI Bemühungen statt, um Vorschläge auszuarbeiten, wie die in MZG verankerten Ideen und Realisierungsvorschläge für eine MZKG adaptiert und modernisiert werden können. Der HCAI-Diskurs hat diverse Modelle für eine MZKG hervorgebracht, die den MZG-Ansatz für die Gestaltung von KI-Systemen erweitern. Erkennbar werden in solchen Modellen die Versuche, eine MZKG möglichst ganzheitlich zu fassen, was beispielsweise durch eine höhere Anzahl von Prozess- und Modellelementen deutlich wird, sowie von den Bestrebungen, technische, organisatorische, soziale und ethische Gestaltungsfaktoren sinnig und realisierbar zu verknüpfen. Resultat dieser Bemühungen sind Modelle, die ein integratives Verständnis von MZKG repräsentieren und im Vergleich mit Modellen zur MZG eine deutlich höhere Komplexität besitzen.

### 3.3.1 Einordnung in KI-Lebenszyklen

Eines dieser MZKG-Modelle (siehe Abb. 14) stammt von einer Gruppe um Wei Xu (Xu, 2019; Xu, Gao, et al., 2023), der in seiner Tätigkeit als Principal Engineer bei Intel und Professor für Human Factors und Engineering Psychology gleichermaßen mit der HCAI-Theorie sowie der Gestaltungspraxis vertraut ist.

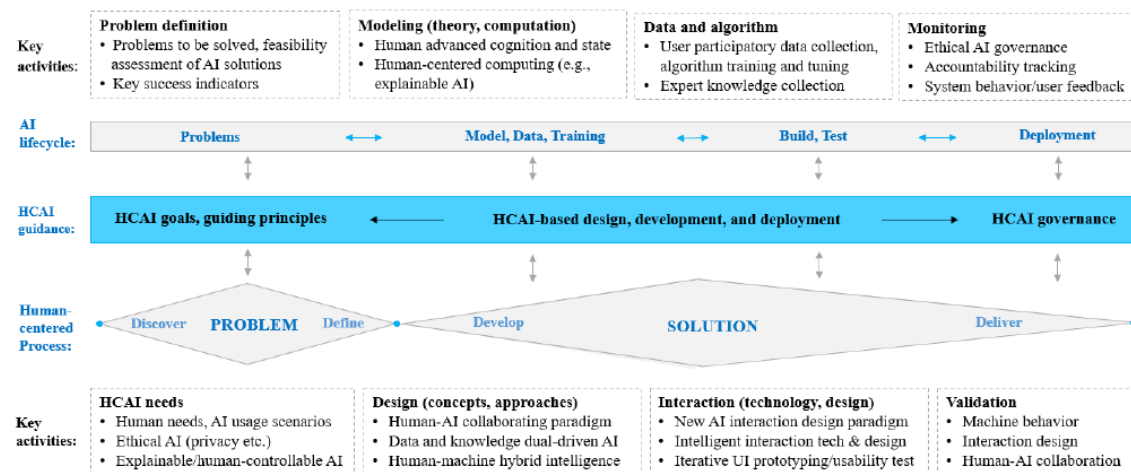


Abbildung 14: MZKG-Modell nach Xu et al. (2023, S. 8)

Motivation für den Entwurf und die Weiterentwicklung des Modells ist es, Gestaltungspraktikern einen möglichst konkreten Vorschlag zu unterbreiten, um die Ideen und Konzepte von HCAI in der Praxis zu realisieren. Das Modell greift MZG in Form des Double-Diamond-Modells von Design auf, einer in HCI etablierten Form von MZG (z.B. Norman, 2013, S. 220). Dieses besteht aus zwei Phasen mit den insgesamt vier Prozessschritten (1) Problemverständnis, (2) Problemdefinition, (3) Lösungsentwicklung und (4) Lösungsüberbringung. Das Modell greift zudem den gesamten Lebenszyklus von KI-Systemen auf, der konkrete Aktivitäten wie Datensammlung, Datenbereinigung, Algorithmustraining oder Wartung von KI-Systemen erforderlich macht. Diese werden im Modell vertikal in die jeweiligen Phasen des Double-Diamond-Modells eingeordnet, ebenso wie diejenigen Aktivitäten, die sich aus der HCAI-Gestaltungsphilosophie ableiten (z.B. Anforderungsdefinition, Usability-Testing). Die HCAI-Gestaltungsphilosophie, die für die Modellentwickler sämtliche Erkenntnisse der MZG-Forschung beinhaltet und repräsentiert, soll den gesamten Gestaltungsprozess anleiten:

*„putting human needs, values, wisdom, capabilities, and roles in the first place in the design, development, and deployment of intelligent systems“* (Xu, Gao, et al., 2023, S. 8).

Damit handelt es sich bei diesem Modell um einen Vorschlag, ein Modell von MZG mit den Modellen zu Lebenszyklen von KI-Systemen zu verknüpfen und die Verknüpfung organisatorischer und inhaltlicher Tätigkeiten auf zeitlicher Perspektive zueinander einzuordnen und in Einklang zu bringen.

### 3.3.2 Information des politischen Geschehens

Ein weiteres, mehr übergreifendes Modell zur MZKG stammt von Jan Auernhammer (Auernhammer, 2020), der sich an der Stanford University unter anderem mit der Theorie und Methodologie der Designforschung auseinandersetzt. Er vertritt die Position, dass die Ideen und Gedanken der humanistischen Designforschung (zu der MZG gezählt werden kann) die fachübergreifende Zusammenarbeit von Entscheidungsträgern des Designs, der TG und der Politik in Fragen der KI-Gestaltung anleiten sollte.

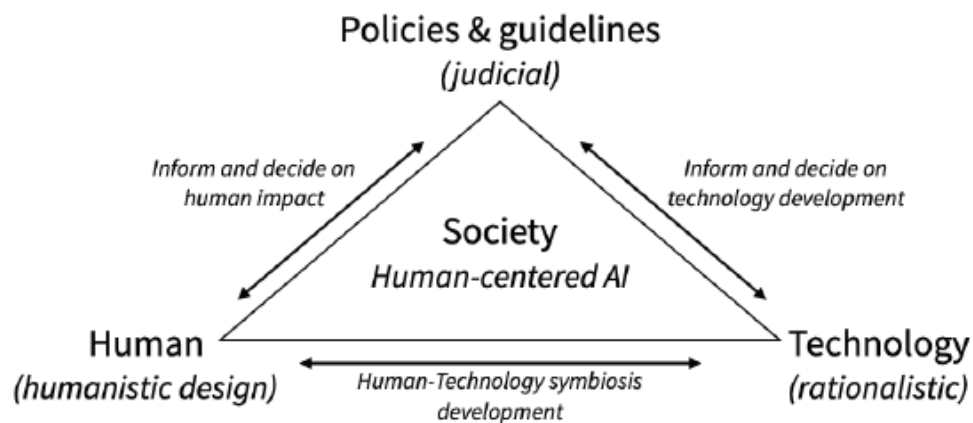


Abbildung 15: MZKG-Modell nach Auernhammer (2020, S. 1328)

In seinem Modell (siehe Abb. 15) ordnet er deshalb die Schule des humanistischen Designs in das Positionsgeflecht aus politischen, technischen und gesellschaftlichen Fragestellungen ein, die bei der Gestaltung von KI-Systemen in HCAI relevant werden. Die humanistische Designforschung hat dabei zur Aufgabe,

*„to examine the potential societal impact of AI technology and systems through “enlightened trial and error” (Auernhammer, 2020, S. 1327).*

Dies soll konkret mithilfe von Experimenten erfolgen, in denen anhand von Prototypen die verschiedenen Auswirkungen von KI-Systemen auf das Wohlbefinden von Individuen sowie der Gesellschaft untersucht werden. Für Jan Auernhammer ist es bei Fragestellungen zur MZKG von essentieller Bedeutung, veraltete Gesetze rasch zu modernisieren, indem in derartigen Experimenten ethische Fragen und nicht-intendierte Folgen von KI-Systemen an konkreten, kontextspezifischen Beispielen untersucht werden. Durch die Verortung in konkreten sozialen Kontexten können zudem technische Fragen der MZKG beleuchtet werden, da die Experimente konkrete Verhältnisse von Menschen mit KI-

Systemen zum Gegenstand haben und Aufschluss darüber geben, ob die menschzentrierten Intentionen von Gestaltern sich in der Praxis auch als solche bewahrheiten. Dieses Modell ist aufschlussreich für den HCAI-Diskurs, da es die verschiedenartigen Fragen und Perspektiven zueinander in Beziehung setzt, mit denen sich die Communitymitglieder in HCAI zunehmend konfrontiert sehen. Das Modell fokussiert dabei besonders die politische Dimension als diejenige, die in die Betrachtungen zur MZKG miteinbezogen werden sollen und die in der Tradition von HCI tendenziell weniger Beachtung finden.

### 3.3.3 Implementierung moderner Werkzeuge

Wieder deutlich stärker auf die Organisation von MZKG-Prozessen ausgerichtet ist das Modell von einer Gruppe um George Margetis (Margetis et al., 2021), welches im Handbook of Human Factors and Ergonomics vorgestellt wird. Dieses Modell nimmt direkten Bezug zum MZG-Modell der Norm DIN EN ISO 9241-210 und erweitert dieses um sechs Konzepte und Werkzeuge, die die Autoren zur Realisierung menschzentrierter KI-Systeme für notwendig und zielführend erachten (siehe Abb. 16).

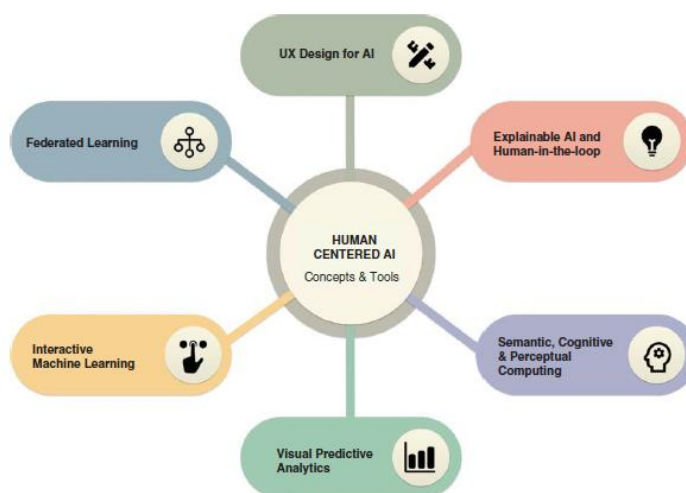


Abbildung 16: Sechs Erweiterungen für das MZG-Modell in DIN EN ISO 9241-210 nach Margetis et al. (2021, S. 1089)

Im Uhrzeigersinn beginnend sieht die erste Erweiterung vor, Erkenntnisse der Forschung zu User Experience (UX; Hassenzahl, 2013; Hassenzahl & Tractinsky, 2006) auf die Gestaltung von KI-Systemen zu übertragen und anzuwenden. Dies erachten die Autoren für notwendig, da viele KI-Systeme aufgrund ihrer Möglichkeit zur Veränderung ihrer Inhalte und Strukturen flexible Interaktionsmöglichkeiten erfordern, die durch traditionelle, mehr deterministische Zugänge zu Interaktionsdesign nicht abgebildet werden können. Die zweite Erweiterung stellt eine Verbindung zu den Erkenntnissen der Felder Explainable AI und Human-AI-Teaming dar. Die Autoren geben wieder, dass menschzentrierte KI-Systeme es Nutzern ermöglichen sollen, nachzuvollziehen, weshalb ein KI-System (k)eine Entscheidung trifft und welche Kriterien es dafür zu Rate zieht. Dies ist notwendig, um die



Vertrauenswürdigkeit der von KI-Systemen getroffenen Entscheidungen beurteilen zu können. Ferner sollen KI-Systeme durch von Menschen getätigte Inputs über „their knowledge about the world and their personal perception“ (Margetis et al., 2021, S. 1091) dazulernen, um so die Interaktionsabläufe kontinuierlich an die Bedarfe von Nutzern anzupassen. Zur Unterstützung dieses im KI-Kontext häufig als Human-in-the-loop bezeichneten Konzepts soll auch die dritte Erweiterung beitragen, welche mit Semantic, Cognitive & Perceptual Computing beschrieben wird. Damit fordern die Autoren, computational models aus sozialwissenschaftlichen Fachbereichen für die Gestaltung von KI-Systemen aufzugreifen. Ähnlich verhält es sich mit Erweiterung Vier (Visual Predictive Analysis), die das Aufgreifen von Modellen und Methoden zur Visualisierung großer Datenmengen beschreibt. Als Interactive Machine Learning beschreiben die Autoren in Erweiterung Fünf das direkte Einbinden von Nutzern in die Trainingszyklen von Machine Learning Anwendungen, sodass nicht die Gestalter der Systeme aufgrund von Nutzerfeedback Veränderungen an den Systemen vornehmen. Die sechste Erweiterung, das Federated Learning, sieht vor, für das Training von Machine Learning Anwendungen verschiedene Endgeräte zu verwenden, die jeweils mit lokalen Datensätzen arbeiten. Hiermit soll der Wahrung der Privatsphäre von Nutzern Rechnung getragen werden, da in traditionellen Trainingsverfahren lediglich mit einem zentralen Datensatz gearbeitet wird. Dieses erweiterte Verständnis von MZKG überführen die Autoren anschließend in ein Modell, das an das Modell der Norm DIN EN ISO 9241-210 angelehnt ist, und konkrete Tätigkeiten zu den sechs Erweiterungen beinhaltet (siehe Abb. 17).

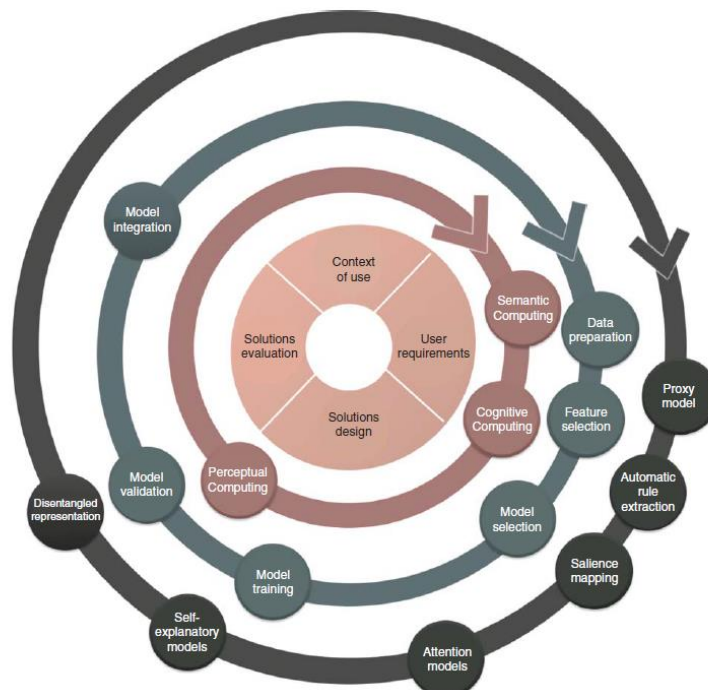


Abbildung 17: MZKG-Modell nach Margetis et al. (2021, S. 1099)

Im Vordergrund dieses Modells zur MZKG steht eindeutig das zyklische Gestaltungsvorgehen, das auch für die MZG von zentraler Bedeutung ist. Im Zentrum stehen die Gestaltungsprinzipien der MZG, die erweitert werden durch die verschiedenen Tätigkeiten, die in KIGP (parallel) durchgeführt werden müssen. Auch dieses Modell veranschaulicht deshalb nachvollziehbar, wie komplex Gestaltungsprozesse der MZKG tatsächlich sind.

### 3.3.4 Balancierung von Autonomie und Automatisierung

Ein weiteres Modell zur MZKG stammt von Ben Shneiderman, Professor für Computer Science am University College of Maryland. Er widmet sich in seinem Buch *Human-centered AI* (Shneiderman, 2022) ebenfalls den konzeptionellen und methodischen Herausforderungen der MZKG und stellt zu Anfang seines Buches die verschiedenen Elemente und Charakteristika dar, die nach seiner Auffassung für die MZKG von Bedeutung sind (siehe Abb. 18). Ein Element stellen verschiedene Konzeptionen menschlichen Lebens und Handelns dar, die Ben Shneiderman unter der Bezeichnung Stakeholders zusammenfasst. In der Mitte der Abbildung sind die verschiedenen Gestaltungsfaktoren dargestellt, sowie die Ziele, die in Prozessen der MZKG angestrebt werden. Bei ihnen sind verschiedene Ausprägungen sozialer, ethischer, rechtlicher, und organisatorischer Faktoren erkennbar, wobei Ben Shneiderman im Vergleich mit anderen Auseinandersetzungen zur MZKG zusätzlich auf die juristischen Gestaltungsfaktoren eingeht. Auch setzt er sich explizit mit Faktoren auseinander, die die Ziele der MZKG gefährden (Threats).

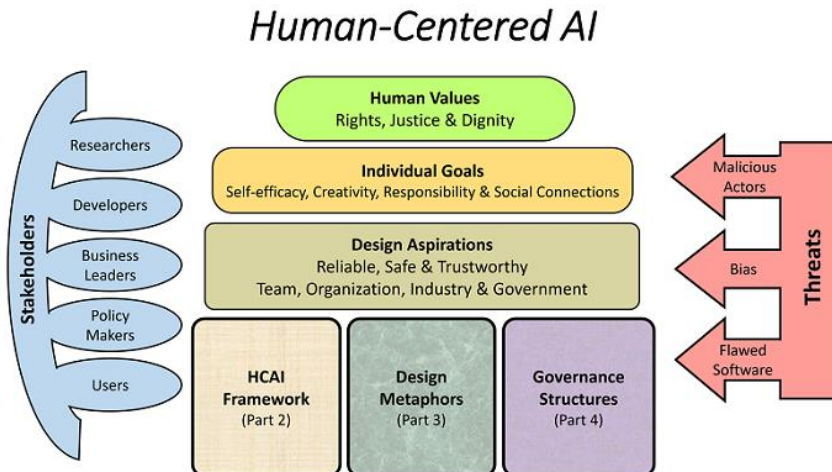


Abbildung 18: Relevante Faktoren einer MZKG nach Shneiderman (2022, S. 28)

Ausgehend von diesem Verständnis von MZKG entwickelt Ben Shneiderman ein zusammenhängendes, mehrdimensionales MZKG-Modell, das sich primär auf den Aspekt der Autonomie und der Handlungsfähigkeit von Menschen in der Interaktion mit KI-Systemen fokussiert und diese den technischen Möglichkeiten zur Automatisierung gegenüberstellt (siehe Abb. 19).



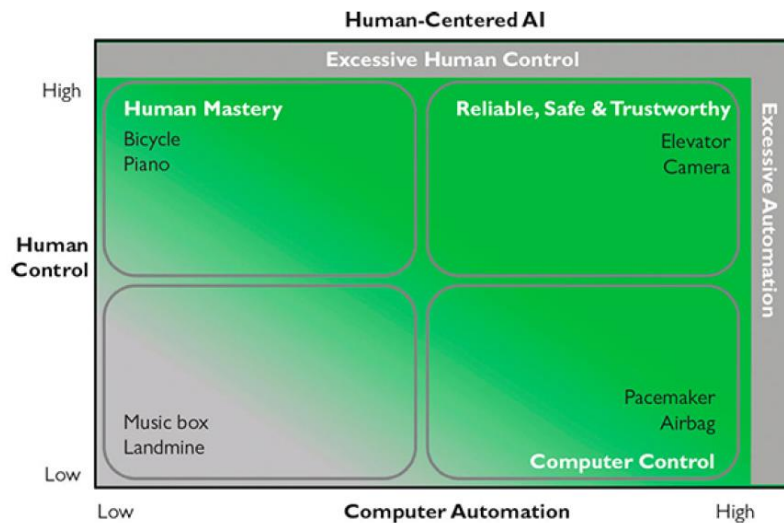


Abbildung 19: MZKG-Modell nach Shneiderman (Shneiderman, 2022, S. 86)

Nach seiner Auffassung hat eine MZKG zur Aufgabe, KI-Systeme zu gestalten, die von Menschen als zuverlässig, sicher und vertrauenswürdig bewertet werden können (rechter oberer Quadrant). Dies kann bei KI-Systemen erzielt werden, indem einerseits die technischen Möglichkeiten zur Automatisierung weit ausgeschöpft werden (x-Achse) und gleichzeitig dafür Rechnung getragen wird, dass der Mensch die nach menschlichem Ermessen zentralen und bedeutsamen Entscheidungen trifft (y-Achse). Zentrale Aufgabe der MZKG ist es demnach, Möglichkeiten für die Herstellung einer angemessenen und zielführenden Balance zwischen menschlicher Autonomie und technischer Automation zu entwickeln. Dabei gilt es darauf zu achten, Extremfälle von menschlicher oder technischer Autonomie abzuwenden (Graubereiche oben, rechts). Dieses Modell ermöglicht es dem HCAI-Diskurs auf Detailebene die verschiedenen Facetten von Autonomie und Automatisierung zu erkennen, zu diskutieren und konkrete Entscheidungsfindungen in Gestaltungsprojekten zu fördern. Ben Shneiderman weist ausdrücklich darauf hin, dass es in Autonomiefragen in der MZKG nicht angebracht ist, eindimensionale Betrachtungsweisen anzuwenden, die die technische Automatisierung auf Kosten menschlicher Autonomie beinhalten.

## 3.4 Trends

Wie aus den Modellen ersichtlich wird, beinhalten die Aufgaben und Zielsetzungen der MZKG eine Vielzahl von Inhalten, die auf verschiedenen Abstraktionsebenen konzipiert werden und in denen verschiedene Formen von Komplexität deutlich werden. Auch werden die Modelle in der HCAI-Community unter teilweise sehr verschiedenen Gesichtspunkten diskutiert, kritisiert und weiterentwickelt. Neben den Modellen zur MZKG sind in HCAI noch weitere Inhalte Gegenstand richtungsweisender Diskurse, die ich nachfolgend schlaglichtartig vorstelle. Dazu greife ich Keynotes auf, die im Zeitraum von 2022 bis 2023 auf einschlägigen Fachtagungen gehalten wurden.

### 3.4.1 Realisierung ko-kreativer und verantwortungsvoller KI

Mary Lou Maher, Professorin am Department of Software Information Systems der University of North Carolina, hat sich in einer Keynote mit den beiden Fragen auseinandergesetzt, wie KI-Systeme gestaltet werden können, um (1) bei Menschen Kreativität und Inspiration zu fördern und um (2) dem Anspruch einer verantwortungsvollen (responsible) KI gerecht werden zu können (Maher, 2022). Derartige KI-Systeme bezeichnet sie als *human-centered co-creative systems*. Um diese Fragen zu beleuchten, gestaltet und evaluiert sie KI-Systeme für Mensch-KI-Kooperationskontexte, die mit Menschen in natürlicher Sprache kommunizieren sollen, um ihr Interaktionsengagement zu fördern (z.B. M. Muller et al., 2024). Dabei begegnen ihr auch Fragestellungen zu ethischen Implikationen ko-kreativer KI-Systeme, zu deren Untersuchung sie auf Erkenntnisse der Pädagogik und Robotik zurückgreift. Sie sieht sich häufig mit der konkreten Frage konfrontiert, welche Werte in der MZKG in den Vordergrund gerückt werden sollen und insbesondere *wessen* Werte besondere Berücksichtigung erfahren sollen. Sie stellt somit auch die Frage, welche Vorstellungen des Menschen in der MZKG zugrundet gelegt werden können oder sollten. Um Antworten auf diese Fragen zu erhalten, setzt sie sich in Studien unter Anwendung von innovativen Methoden wie dem Design Fiction mit den Wertauffassungen von verschiedenen Stakeholdergruppen auseinander und befragt sie zu ihren Auffassungen der Bedeutung von verantwortungsvoller KI (siehe Abb. 20).

## End User Perception of Responsible AI

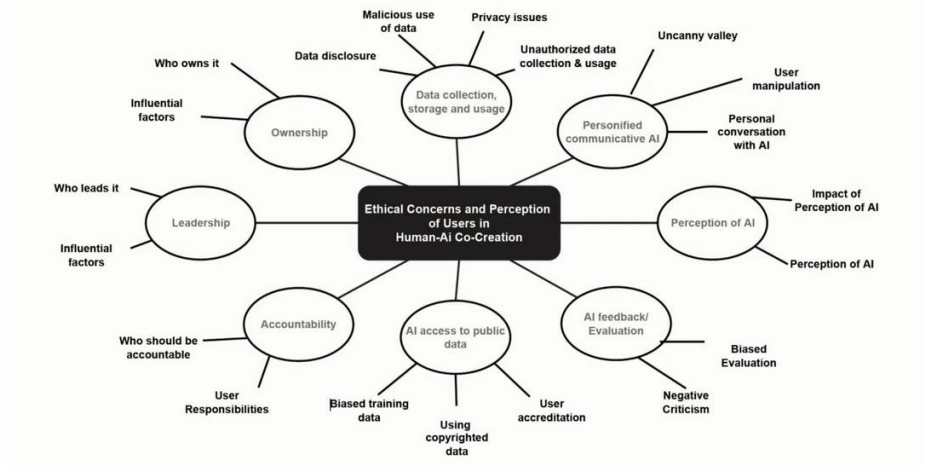


Abbildung 20: Nutzerauffassungen zu Responsible AI (Maher, 2022)

### 3.4.2 Reflektierter Umgang mit Konzeptionen zu menschlichem Leben und Handeln

Mit der Bewältigung ethischer Herausforderungen der MZKG setzt sich auch Fernanda Viégas in ihren Funktionen als Professorin für Computer Science der Harvard School of Engineering sowie als Google-Researcherin auseinander (Viégas, 2022). Auch ihr begegnet bei ihrer Arbeit häufig die Frage, welche Menschen das „human-centered“ in HCAI adressiert, da die Antworten auf diese Frage elementare Wegweiser für die Gestaltung von Interaktionsschnittstellen und Interfaces für KI-Systeme sind. Sie weist deshalb auf ein Spektrum von Stakeholdern hin, das bei der Realisierung von Transparenz, Erklärbarkeit und Kontrollierbarkeit von KI-Systemen Berücksichtigung erfahren sollte (siehe Abb. 21).



Abbildung 21: Spektrum von Stakeholdern in der KI-Gestaltung (Viégas, 2022)

Es sei elementar, sich in einem Gestaltungsprojekt vor Augen zu führen, wie die relevanten Stakeholdergruppen zusammengesetzt sind – Experten einer Domäne stellen an KI-Systeme andere Anforderungen wie beispielsweise Gelegenheitsnutzer. Fernanda Viégas weiß aus ihrer Praxiserfahrung auch um die konkreten Problemstellungen, die dieser Umstand für KI-Community hervorruft. Sie beschreibt, dass es der KI-Community nicht gelänge, Nutzer in die Gestaltung und Evaluation von KI-Systemen zu integrieren und sie von der Genauigkeit der in KI-Systemen verwendeten Machine Learning Modelle zu überzeugen. Sie betrachtet das Problem in ihrer Keynote genauer und stellt dar, weshalb es sich bei der Bewältigung dieser Herausforderung in doppelter Hinsicht um eine kritische Zielsetzung für die MZKG handelt. Eine Lösungsmöglichkeit illustriert sie an einer Use Case-Studie im Medizinbereich, in der sie in Kooperation mit medizinischem Personal und der Anwendung der Prototyping-Methode ein Interface für eine KI-Anwendung entwickelte. Dieses UI ermöglichte es Nutzern, ihre Bewertungskriterien auf die Auswertung medizinischer Bilddaten anzuwenden, ohne ausschließlich auf die stellenweise intransparenten automatisierten Bewertungskriterien des Systems zu vertrauen. Die Studienteilnehmer konnten durch das alternative Interface ein höheres Vertrauen zum System und den von ihm gelieferten Resultaten aufbauen.

#### 3.4.3 Zur Bedeutung der Humanities für HCAI

Eine Verbindung zwischen der HCAI-Community und der Community der (Digital) Humanities stellt Lauren Klein in einer Keynote her (Klein, 2022). Sie ist Professorin für Quantitative Theorie und Methoden, sowie für Englisch an der Emory University in Atlanta. Ihre Interpretation von HCAI besagt, dass menschenzentrierte KI-Systeme zum Wohle der Menschheit (humanity) gestaltet werden sollten. Daraus leitet sie ab, dass dies die Analyse der kulturellen Aspekte von KI-Gestaltung bedingt, wie sie in den Traditionen der Humanities erfolgt – unter der Analyse und der Berücksichtigung von Kontext. Besonders bedeutsam ist für Lauren Klein hervorzuheben, dass es sich bei KI-Systemen um sozio-technische Systeme handelt, deren Gestaltungsprozesse kulturelle Vorgänge darstellen, weshalb auch KI-Systeme kulturelle Erzeugnisse darstellen. Diese Auslegung erläutert sie anhand von verschiedenen Beispielen, wovon eines die in KI-Systemen verwendeten Datensätze darstellen. Solche Datensätze werden in den Humanities als „objects of culture“ betrachtet, in denen beispielsweise bestimmte Machtverhältnisse berücksichtigt oder ignoriert werden oder lediglich ausgewählte Erfahrungswerte repräsentiert werden. HCAI habe deshalb zur Aufgabe, die Vorgänge von KI-Gestaltung als constructions of culture aufmerksam zu reflektieren und in ihrer Bewertung Machtverhältnisse sowie soziale, historische, geographische und weitere Kontextfaktoren miteinzubeziehen. Hierfür, so die Position von Lauren Klein, benötigt die HCAI-Community die Erfahrungswerte der Humanities.

### 3.4.4 Verantwortungsbereiche von HCAI

Peter Norvig, Education Fellow des Stanford Institute for Human-Centered Artificial Intelligence und Engineering Director bei Google, gibt in einer Keynote seine Perspektive auf das Feld HCAI wieder (Institute for Experiential AI, 2023). Dafür greift er die Entstehung von KI aus dem Software Engineering auf und zeigt, dass in HCAI ein ganzheitlicher Ansatz verfolgt wird, um die von Menschen angestrebten Ziele durch KI-Systeme technisch zu unterstützen. Nach seiner praktischen Erfahrung hat sich in der KI-Gestaltung erwiesen, dass es bei der **Identifikation und Realisierung „tatsächlicher Ziele“ von Menschen** häufig nicht um Fragen und Aufgaben mathematischen Typs handelt, sondern um Wertefragen, zu deren Beantwortung Erkenntnisse der (Technik-)Ethik hilfreich sind. Er fährt fort, das Feld HCAI zu charakterisieren und verwendet dafür verschiedene Charakteristika, die deutlich die in der MZG verankerten Gedanken wiedererkennen lassen (siehe Abb. 22).



Abbildung 22: Charakterisierung von HCAI nach Norvig (Institute for Experiential AI, 2023)

Besondere Bedeutung erfährt für Peter Norvig dabei das Element der Multidisziplinarität. Dieses sei in HCAI von zentraler Bedeutung, um beispielsweise mehr Use Cases für KI und Gefahren von KI identifizieren zu können, Biases in KIGP zu reduzieren und Zugriff auf verschiedene Zugänge und Methoden zur Gestaltung zu erlangen. Diese seien in der KI-Gestaltung umso mehr von Bedeutung, um die technischen, gesellschaftlichen und wirtschaftlichen Ziele von KIGP in Einklang zu bringen.

## 3.5 Zusammenfassung

Im Feld HCAI werden KI-Gestaltung und gesellschaftliche Bedarfe zusammenhängend betrachtet und vor diesem Hintergrund nach Lösungen gesucht, um KI-Systeme mensch-zentriert zu gestalten. Diese Zielsetzung erfordert einerseits eine angemessene Ausein-andersetzung mit den verschiedenen relevanten technik-, sozial- und geisteswissenschaftli-chen Fragestellungen, andererseits jedoch auch die Priorisierung des Gestaltungs-pragmatismus. Beide Aufgaben in Einklang zu bringen haben sich verschiedene wissen-schaftliche Einrichtungen und Kooperationen in HCAI zum Ziel gesetzt, die auch Anknüp-fungspunkte zu Feldern wie dem Ethical AI und dem Explainable AI beinhalten. Zudem ist in den Debatten zur MZKG der Bezug zu Erkenntnissen wie der Psychologie oder der Organisationforschung recht hoch, was die verschiedenen Debatten und Publikationen zu Stichworten wie dem Human-AI-Teaming und dem Human-in-the-loop zeigen. Initiativen wie das Humane AI Net der EU oder die Gründung ganzer Universitätsfakultäten zeigen, mit welcher Ernsthaftigkeit an den Themen der MZKG gearbeitet wird und welche Dyna-mik im Feld HCAI in 2024 vorhanden ist. Zwei Schwerpunktaspekte des HCAI-Diskurses beschäftigen sich dabei (1) mit den verschiedenen Konzeptionen menschlichen Lebens und Handelns, die in Gestaltungsprozessen zum Tragen kommen können und (2) der Entwicklung von Modellen und Methoden, die eine systematische Gestaltung und Evalua-tion menschzentrierter KI-Systeme ermöglichen.

In diesen Debatten werden punktuell Bezüge zu TA sichtbar, wenn beispielsweise Fragen zu Gestaltungsintentionen, der Bedeutung von Partizipation in Gestaltungsvorgängen oder der Einbeziehung technikethischer Erkenntnisse diskutiert werden. Ein Hinterfragen und Reflektieren über das Zustandekommen von Technik ist regelmäßig Gegenstand von TA, genauso wie das Bestreben, durch die Beteiligung an Diskursen zur Schaffung von Technik zu gesellschaftlich wünschenswerten Effekten beizutragen. Somit stellt sich die Frage, ob Erkenntnisse aus TA anschlussfähig sind, um Beiträge zu den Debatten zu Kon-zeptionen menschlichen Lebens und Handelns und dem Bedarf an Gestaltungsmethoden in HCAI zu leisten. Der Auseinandersetzung mit dieser Fragestellung sind die folgenden Kapitel gewidmet, beginnend mit der Betrachtung von Auffassungen zu TG in TA.

## 4 Gestaltungsvermögen in Technikfolgenabschätzung

In Kapitel 4 vollziehe ich einen Perspektivwechsel hin zu Auffassungen von Technikgestaltung (TG) in Technikfolgenabschätzung (TA). Die Positionen zu TG in TA greifen traditionell Perspektiven der Techniksoziologie und Technikphilosophie auf, weshalb in TA-Arbeiten im Vergleich mit HCI und HCAI andere Ausgangspunkte gewählt werden, wenn Perspektiven auf und Aktivitäten in Gestaltungsprozessen diskutiert werden. Dennoch sind in beiden Communities unverkennbare Gemeinsamkeiten in der Debatte zur KI-Gestaltung vorhanden, da auch in TA die Förderung des gesellschaftlichen Miteinanders eine zentrale Priorität darstellt. Zur Adressierung der verschiedenen sozialen und ethischen Fragestellungen der MZKG verfügt TA über verschiedene Gestaltungsansätze, die eine potenzielle Anschlussfähigkeit zum HCAI-Diskurs aufweisen. Zu diesen Gestaltungsansätzen zählen das Constructive Technology Assessment (CTA) und das Value Sensitive Design (VSD), die der Dissertation als Untersuchungsgegenstand dienen. In Kapitel 4.2 (CTA) und Kapitel 4.3 (VSD) stelle ich die zentralen Annahmen sowie die Methodik beider Ansätze vor. Anschließend diskutiere ich in Kapitel 4.4 diverse Aspekte, die für die Fragestellung der Anschlussfähigkeit zu HCAI bedeutsam sind.

### 4.1 Einführung

Zentraler Gegenstand in TA ist die Analyse von Chancen und Risiken von Technologien und technischen Systemen, die das gesellschaftliche Leben in der Gegenwart bereits in hohem Maße prägen oder es zukünftig mit hoher Wahrscheinlichkeit prägen werden. Maßgebliche Motivation für TA-Aktivitäten ist die aus der Technikgeschichte stammende Erkenntnis, dass von Menschenhand geschaffene technische Erzeugnisse selten ausschließlich diejenigen Zwecke erfüllen, die der Mensch im Vorfeld der Gestaltung identifiziert hat bzw. durch seine Gestaltungsaktivitäten zu realisieren sucht (Jonas, 2019; Ogburn, 1922). Beispiele wie etwa Asbest oder Contergan dienen in TA dazu, die in der Vergangenheit auf tragische Weise zum Vorschein getretenen negativen Folgen technischen Fortschritts zu erläutern und die Notwendigkeit für vorausschauende TA-Arbeit präventiven Charakters aufzuzeigen. Die Wiederholung solcher Geschehnisse abzuwenden und Anzeichen für ihre Entstehung möglichst frühzeitig zu erkennen, stellt ein Hauptanliegen von TA-Arbeit dar (Vorsorgeprinzip; von Gleich, 2013). In TA wird angestrebt, Technik und TG-Prozesse mehr ganzheitlich zu betrachten, um durch das auf diese Weise entstehende vollständigere Bild möglichst sämtliche Folgen für das gesellschaftliche Leben erfassen und antizipieren zu können. Zusätzlich zu den unmittelbar betroffenen Lebensbereichen werden besonders die Lebensbereiche in TA-Untersuchungen miteingeschlossen, die mittelbar durch die Einführung, Nutzung oder Adaption technischer Systeme Veränderungen unterliegen. Entsprechend wird in TA zwischen intendierten und nicht-intendierten Technikfolgen unterschieden (Gloede, 2007), wobei insbesondere den nicht-intendierten

Folgen hohe Aufmerksamkeit gewidmet wird. Die Antizipation und Beurteilung von Technikfolgen macht die Auseinandersetzung mit der Gestaltung einer betreffenden Technik unumgänglich (Grunwald, 2003). Die Ergründung der Gestaltungsintentionen und den dafür eingesetzten Methoden sind regelmäßig Gegenstand von TA-Arbeit, da sie wertvolle Informationen und Anhaltspunkte für die Folgenabschätzung bereitstellen. So erfordert TA-Arbeit gelegentlich, die für TA typische distanzwahrende und auf abstraktem Niveau verbleibende Position zu verlassen, um Folgenwissen aktiv in Gestaltungsprozesse einzubringen und Präventions- und Reflexionsaktivitäten zu motivieren (siehe Abb. 23).

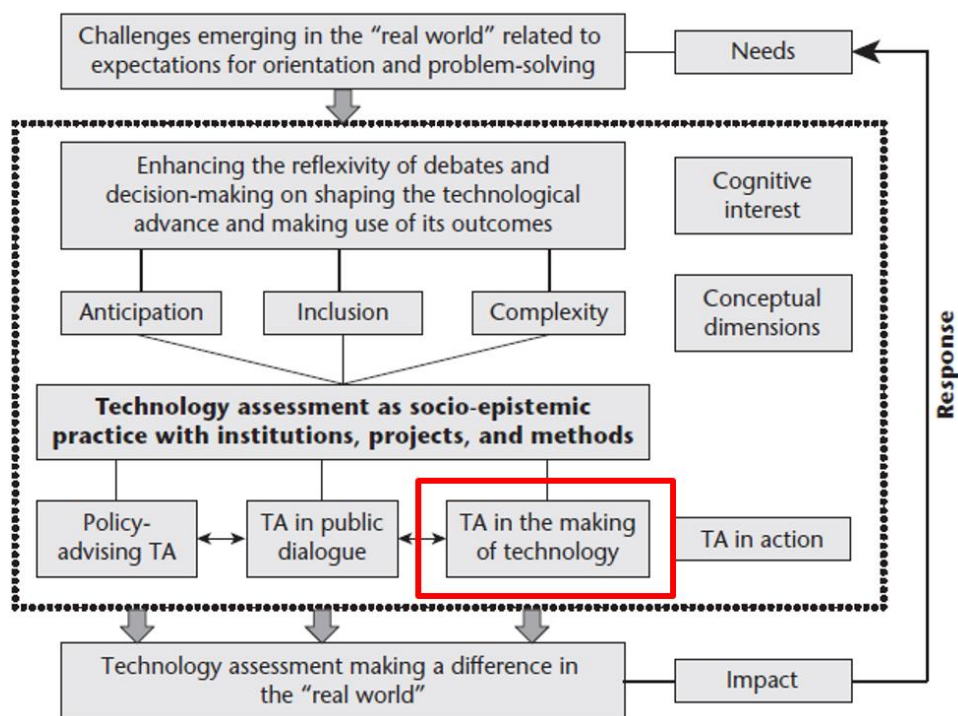


Abbildung 23: TA in the making of technology als ein Element einer TA in action (Grunwald, 2019, S. 89)

Um dieser Aufgabe gerecht zu werden, können in TA-Projekten prinzipiell verschiedene Ansätze gewählt werden. Bei der Auswahl eines Ansatzes werden üblicherweise verschiedene Faktoren berücksichtigt, wie etwa die Charakteristika der betrachteten Technologie, die vorhandenen Stakeholderkonstellationen sowie die konkreten Ziele der TA-Intervention. Unterstützt und systematisiert werden diese Erwägungen in Ansätzen wie dem CTA und dem VSD, wobei besonders CTA aufgrund seiner Entwicklung in der TA-Community ein hohes Ansehen innerhalb der TA genießt.



## 4.2 Constructive Technology Assessment

In CTA wird danach gestrebt, bereits in frühen Phasen von Gestaltungsprozessen die Reflexion über gesellschaftliche Auswirkungen anzustoßen, die mit einer betrachteten Technologie zusammenhängen. Unter Anwendung und Operationalisierung von sozial-konstruktivistischen Positionen und Konzepten soll eine stärkere Beteiligung von sozialen Akteuren in den Gestaltungsprozess erfolgen und die direkte Rückführung von Reflexions-erkenntnissen in Gestaltungsprozesse ermöglicht werden. Es sollen technische Systeme „with desired positive impacts and with few (or at least manageable) negative impacts“ (Rip et al., 1995, S. 3) gestaltet werden, um dadurch Beiträge zu einer „better technology in a better society“ (Schot & Rip, 1997, S. 266) zu leisten.

Das CTA wurde während der 1980er- und 1990er-Jahre in den Niederlanden entwickelt, nachdem das dortige Ministerium für Bildung und Wissenschaft durch ein Memorandum eine Neuausrichtung der niederländischen TA vornahm (Ministerie van Onderwijs en Wetenschappen, 1984, zitiert nach Schot & Rip, 1997). Nach Vorbild des US-amerikanischen Office of Technology Assessment der 1980er-Jahre soll TA auch in den Niederlanden zu einem „broader decision making about science and technology in society“ (Schot & Rip, 1997, S. 252) beitragen. Konkrete Maßnahmen sehen ein „broadening [of] technology development by including more aspects and involving more actors“ (Rip & Robinson, 2013, S. 38) vor, bei dem besonders soziale Aspekte als zusätzliche Gestaltungskriterien etabliert werden. In die Ausarbeitung des Memorandums ist auch der Professor of Philosophy of Science and Technology Arie Rip involviert, der in der darauffolgenden Zeit maßgeblich auf die Ausarbeitung und Weiterentwicklung von CTA hinwirkt (z.B. Rip, 2002, 2008; Rip et al., 1995). Die Arbeit an der CTA-Theorie beginnt daraufhin an der 1986 neu gegründeten Netherlands Organization for Technology Assessment (seit 1994 Rathenau Instituut), der bis heute bedeutendsten TA-Institution der Niederlande. Parallel zur Entwicklung eines ersten ganzheitlichen CTA-Ansatzes erfährt das Konzept einer konstruktiven TA zunehmend internationale Resonanz (z.B. in Remmen, 1995; Tatum, 1997). Mit Beginn der 2000er-Jahre wird der CTA-Ansatz verstärkt in Kontexten von Nanotechnologie erprobt und geschärft (Rip, 2008; Robinson, 2010; van Merkerk, 2007). Seitdem haben sich die Anwendungsfelder von CTA stark erweitert, beispielsweise wird es in der Medizininformatik (Høstgaard et al., 2017), im Zusammenhang mit Energiespeichertechnologien (Versteeg et al., 2017) oder nachhaltiger Produktentwicklung (Kuhlmann, 2012) aufgegriffen. In den vergangenen Jahren wird es auch gelegentlich mit KI-Anwendungen in Verbindung gebracht (Schepis et al., 2019; van de Poel, 2020; Xue & Pang, 2022), allerdings überwiegend, um auf theoretischer Ebene Gedanken anzustellen.

### 4.2.1 Zentrale Annahmen und Konzepte

Die Theorie von CTA fußt auf dem sozialkonstruktivistischen Diskurs, der mit Beginn der 1980er-Jahre in den Feldern der Techniksoziologie und Science and Technology Studies (STS) Einzug hält (Bijker et al., 1987; D. A. MacKenzie & Wajcman, 1999; Pinch & Bijker, 1984). Technische Innovationen werden darin als soziales Konstrukt betrachtet, als ein Produkt sozialer Prozesse, die durch menschliche Aushandlungen und Entscheidungen zustande kommen. Dies stellt eine klare Gegenposition zum damals in der sozialwissenschaftlichen Technikforschung überwiegend vertretenen Position des Technikdeterminismus (z.B. Lutz, 1987; Smith & Marx, 1994) dar, der Technik als maßgeblichen Treiber für soziale, politische und kulturelle Entwicklungen betrachtet. Im technikdeterministischen Denken wird der Gesellschaft wenig bis gar kein Spielraum für eine (Mit-)Gestaltung von Technik beschieden, weshalb sie sich technischen Innovationen zeitlich nachgelagert und überwiegend machtlos anzupassen hat. Der Sozialkonstruktivismus stellt diese Grundposition in Frage, weshalb er heute für einen Paradigmenwechsel in der sozialwissenschaftlichen Technikforschung steht.

CTA greift die sozialkonstruktivistische Bewegung auf verschiedene Weise auf, was bereits in der Bezeichnung als Constructive Technology Assessment deutlich wird. Die Bedeutung des Konstruktiven ist in CTA mehrfach belegt. So kann damit (1) die Konstruktion von Wissen im Sinne eines gesellschaftlichen Lernvorgangs „to improve public understanding of science and technology“ (Rip et al., 1995, S. 6) gemeint sein, (2) die „actual construction“ (Rip, 2008, S. 148) von Technik, oder auch (3) das Ziel, einen konstruktiven im Sinne von fördernden Beitrag zum Geschehen von Gestaltungsprozessen und der gesellschaftlichen Einbettung von Technik zu leisten (Konrad, 2021). Mit Blick auf die angestammte TA-Community versteht sich CTA darüber hinaus (4) als Vorschlag zu einem „constructive way of doing technology assessment“ (Rip, 2018, S. 97). Auf sprachlicher Ebene wird in CTA demnach versucht, den positiv assoziierten Interpretationsspielraum des (sozial-)konstruktiven Bedeutungsraums maximal auszuschöpfen.

Ein weiteres zentrales Element für CTA ist der für den Sozialkonstruktivismus zentrale Ansatz des Social Construction of Technology (SCOT; Bijker et al., 1987). Darin wird die Gestaltung technischer Artefakte als evolutionärer Vorgang konzipiert, den Menschen durch das permanente Treffen von Entscheidungen kontinuierlich vorantreiben. In dieser Vorstellung berücksichtigen Menschen beim Treffen ihrer Entscheidungen sämtliche kulturellen, sozialen, politischen, ethischen und weiteren für sie relevanten Aspekte. Diese Vorstellung von TG als einem evolutionären, entscheidungsbasierten und stetig voranschreitenden Prozess übernimmt CTA, und erweitert sie hinsichtlich dem Aspekt, dass technische Artefakte sich auch auf die Entscheidungsfindungen von Menschen aus- und rückwirken. In CTA werden Gestaltungsprozesse als differenzierte Prozesse soziotechnischer Wechselwirkungen verstanden, als eine „co-evolution of science, technology and society“ (Rip, 2002, 2008), in der sich Gesellschaft und Technik in ihrer Entwicklung wechselseitig beeinflussen (co-production of impacts) und deshalb nicht voneinander getrennt betrachtet werden sollten (Konrad, 2021, S. 214). CTA kritisiert damit den in den 1990er- und 2000er-Jahren in vielen Gesellschaften üblichen two-track-approach, der

eine Verantwortungsseparation des gesellschaftlichen Technologiemanagements auf die Gestaltung von Technik („technology promotion“) und die Regulierung von Technik („technology control“) vornimmt. Die gedankliche und letztlich institutionelle Trennung dieser beiden Seiten habe in vielen Gesellschaften eine ungleiche Machtverteilung bewirkt:

*„Thus, in our modern societies there is an asymmetry between ‘impactors’ (those at the source of technological impacts) and ‘impactees’. This can be a difference in power, but is always a difference in timing. Initiators of technological development know more, and have invested more, at an early stage, and impactees and spokespersons for society have to wait and in a sense, follow their lead”* (Rip, 2008, S. 146).

Diesen Zustand möchte CTA überwinden, indem es „bridges between promotion and control activities“ (Rip et al., 1995, S. 3) schafft. Konkret soll ein solcher Brückenschlag durch die mit der CTA-Methodik vorgeschlagenen soft interventions erfolgen, durch die die TA-Perspektive durch punktuelle Interventionen in Gestaltungsprozessen platziert werden kann (Rip & te Kulve, 2008). Die damit verbundene Zielsetzung ist, dass bereits frühzeitig negative Folgen von Technik (und ihrer Gestaltung) sichtbar werden, der jedoch aufgrund der noch weitgehend gestalt- und wandelbaren Technik entgegengewirkt werden kann. Auf diese Weise kann CTA auch dem in TA-Arbeiten häufig angetroffenen Collingridge-Dilemma (Collingridge, 1980) begegnen, da es Erkenntnisse von TA-Arbeit frühzeitig in Gestaltungsprozesse zurückführt:

*„CTA wants to intervene/modulate at an early stage, with the expectation that things will go better in later stages, somehow”* (Rip, 2018, S. 110).

Die Wortwahl der Intervention, die an dieser Stelle gewählt wird, ist bezeichnend für den in CTA verankerten, traditionellen TA-Charakter – Intervention und Modulation von Gestaltungsprozessen zielen mehr auf die Veränderung von Bestehendem ab, wie auf die für viele Technik- und Ingenieurwissenschaften typischen Zielsetzung der gänzlichen Neuschaffung von Artefakten, Systemen und Prozessen.

#### 4.2.2 Vorgeschlagene Methodik

Die zentralen Annahmen und Konzepte werden in CTA in eine Methodik überführt, die es ermöglichen soll, TA-Aktivitäten in Gestaltungsprozesse einzubringen. Grundsätzlich lautet der Anspruch, die CTA-Methodik auf Gestaltungsprozesse *sämtlicher* Technologien anwenden zu können – es besteht somit kein Fokus auf eine bestimmte Technologieform, wie etwa KI. Im Verständnis in CTA findet TG in „Technologiefeldern“ statt, in welchen verschiedene Akteure aktiv sind und die durch ihre Handlungen und Entscheidungen eine Technologie oder ein bestimmtes technisches Artefakt gestalten. Diese Akteure werden in CTA beabsichtigt, bei der Artikulation ihrer Positionen und Strategien zu unterstützen und somit die Reflexion im Gestaltungsprozess zu fördern (Rip, 2008, S. 147). Die damit verbundene Annahme lautet, dass eine bewusste und reflektierte Betrachtung der eigenen Strategien sowie die der weiteren beteiligten Akteure im Technologie-

feld nicht nur für ein besseres Verständnis der Dynamiken im Technologiefeld sorgt, sondern auch frühzeitig für die Gesellschaft potenziell problematische Entwicklungen aufzeigen kann. Um die Reflexion noch weiter zu fördern, sieht die CTA-Methodik die Beteiligung sogenannter *sozialer* Akteure in Gestaltungsprozessen vor. Soziale Akteure bezeichnen in diesem Zusammenhang primär soziale Gruppen, die wirtschaftliche, gesundheitliche, kulturelle oder anderweitige Auswirkungen der betreffenden Technologie oder des betreffenden technischen Systems erfahren werden, jedoch bislang nicht direkt am Gestaltungsprozess beteiligt sind (stellenweise auch als „*impactees*“ bezeichnet). Durch die Integration sozialer Akteure soll die inhaltliche sowie personelle Erweiterung von Gestaltungsprozessen erfolgen, da soziale Akteure durch das Ein- und Vorbringen ihrer Standpunkte die Aufmerksamkeit in Gestaltungsprozessen vornehmlich auf nicht-technische Aspekte richten. Als Mittel zur Reflexions- und Dialogförderung aller beteiligten Akteure sieht die CTA-Methodik die Entwicklung von soziotechnischen Szenarien vor. Diese sollen verschiedene zukünftige Entwicklungen des Technologiefelds skizzieren, die aufgrund von beispielsweise Handlungen der Akteure im Feld oder auch durch feldexterne Entwicklungen vorstellbar sind. Die soziotechnischen Szenarien fungieren wiederum als Einstiegspunkt und Diskussionsgrundlage für Workshops, an denen sowohl Feldakteure wie auch soziale Akteure teilnehmen und ihre verschiedenen Positionen und Strategien austauschen und artikulieren sollen. Auf diese Weise sollen allen Akteuren Einblicke in die Positionen und Argumentationsstrategien der jeweils beteiligten Akteure im Sinne eines „*probing [of] each other's worlds*“ (Rip & te Kulve, 2008) ermöglicht werden. Angeleitet werden die teilnehmenden Akteure in den Workshops durch moderierende CTA agents, die typischerweise Expertise in Techniksoziologie, TA oder einer verwandten Fachrichtung besitzen.

In Summe kann die CTA-Methodik somit in drei Aktivitäten unterteilt werden: (1) die Analyse des Technologiefelds, (2) der Entwicklung soziotechnischer Szenarien und (3) der Durchführung von CTA-Workshops zur Reflexion der im Technologiefeld vorhandenen Strategien (siehe Abb. 24).

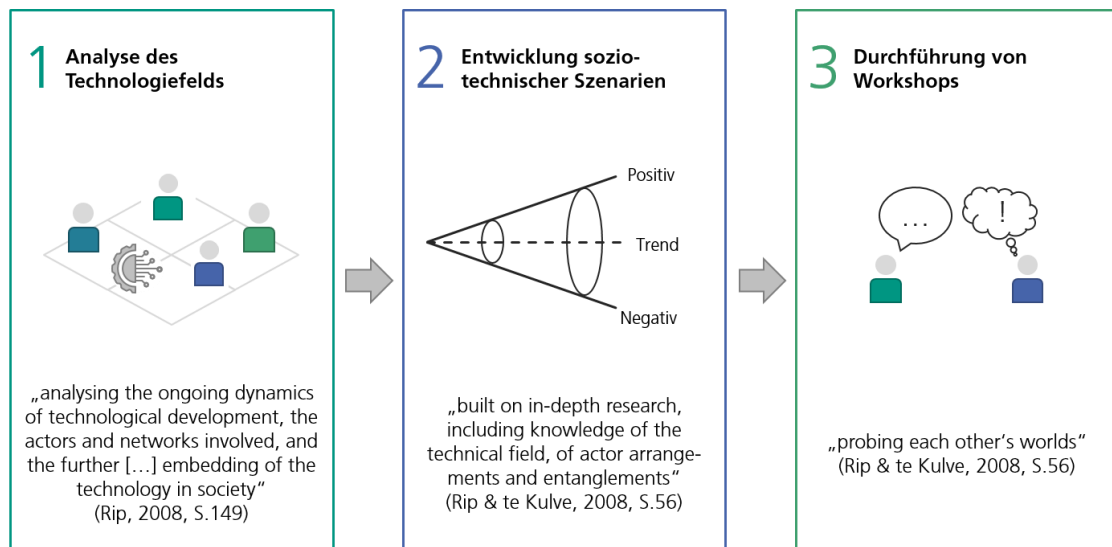


Abbildung 24: Die drei Aktivitäten der CTA-Methodik

Im Hinblick auf zukünftige Gestaltungsvorgänge lautet die Annahme in CTA, dass sämtliche beteiligten Akteure durch die reflektierenden Betätigungen sich über mögliche negative gesellschaftliche Auswirkungen ihres Handelns bewusst werden und künftig Maßnahmen ergreifen, um diese abzuwenden oder um alternative Gestaltungswege zu finden. Auf diese Weise soll insbesondere bei den gestaltenden Akteuren eine gewissermaßen intrinsische Motivation angeregt werden, um so aus gesellschaftlicher Perspektive „bessere“ Technologien zu gestalten.

### Analyse des Technologiefelds

Für die Analyse des Technologiefelds ist in CTA vorgesehen, dass sich CTA agents in das entsprechende Technologiefeld hineinbegeben, in dem eine Technologie gestaltet wird. Ziel dieser Aktivität ist, die im Feld aktiven Akteure, den aktuellen bzw. vorgesehenen Entwicklungszustand der Technologie, sowie ihre angedachte gesellschaftliche Einbettung zu ermitteln. Das Hineinbegeben in das Technologiefeld kann auf unterschiedliche Weise erfolgen, typischerweise sind Besuche von verschiedenen Einrichtungen und Veranstaltungen erfolgsversprechend:

*„research laboratories, conferences, workshops, agenda setting and planning meetings, roadmapping events, public debates anticipating on issues related to technology developments“ (Rip & Robinson, 2013, S. 38).*

Von zentralem Interesse sind dabei die sich abzeichnenden Muster und Entwicklungen auf der soziotechnischen Ebene, da sie die Zusammenhänge der Gestaltung und der zukünftigen gesellschaftlichen Einbettung der betrachteten Technologie offenbaren und auch erste Hinweise auf tatsächliche und potenzielle Konfliktfelder eröffnen. Geleitet werden soll die Analyse des Technologiefelds von den theoretischen Interessen der CTA

agents, die diese im Vorfeld formulieren. Da CTA auf die Gestaltung sämtlicher Technologien angewandt werden können soll, kann der inhaltliche Fokus der Technologiefeldanalyse ein breites Spektrum umfassen. So kann der Fokus beispielsweise auf der Technologie als solche liegen, es können aber auch spezifische technische Produkte oder Verfahren der betreffenden Technologie im Vordergrund stehen. Für die Durchführung der Technologiefeldanalyse können die CTA agents prinzipiell sämtliche Methoden der sozialwissenschaftlichen Technikforschung anwenden, sofern sie sie bei der Beleuchtung ihrer theoretischen Interessen und der Klärung ihrer Fragen unterstützen. Im Ergebnis sollen die CTA agents einen Überblick über die im Technologiefeld aktiven Akteure besitzen und bezogen auf den Stand der Technik ein ausreichend vollständiges Grundverständnis der technischen Zusammenhänge und Herausforderungen erarbeiten (siehe Abb. 25).

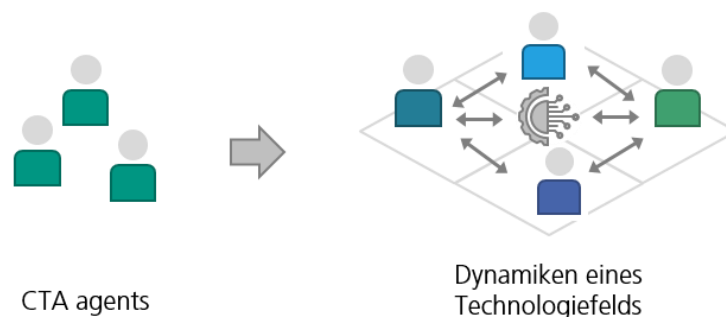


Abbildung 25: In der Technologiefeldanalyse begeben sich CTA agents in ein Technologiefeld hinein

## Entwicklung soziotechnischer Szenarien

Dieses Wissen benötigen CTA agents, um anschließend Szenarien entwickeln zu können, die potenziell bedeutsame soziotechnische Entwicklungen des Technologiefelds beschreiben. Dieser Tätigkeit liegt die Annahme zugrunde, dass mögliche Zukünfte des Technologiefelds bereits in seiner Gegenwart eingebettet sind und aus dem Feld ausgebildet werden (endogenous futures; Moretto et al., 2014). Entsprechend sollen die soziotechnischen Szenarien zukünftige Zustände des Technologiefelds beschreiben, die sich auf bereits stattfindende oder sich abzeichnende Entwicklungen (Konrad, 2021) im Technologiefeld beziehen und diese in bewusst gewählte, positive oder negative Richtungen fortführen. Dieser Realitätsbezug ist von entscheidender Bedeutung für die soziotechnischen Szenarien:

*„In Constructive TA, socio-technical scenarios are not just creative exercises showing possible futures (which is one function of scenarios). They embody and further articulate emerging patterns in interactions, up to paths that actors tend to follow“ (Rip & te Kulve, 2008, S. 50).*

Die konkreten Inhalte und Ausrichtungen der Szenarien sind abhängig von den Erkenntnissen der Technologiefeldanalyse und können sich inhaltlich auf verschiedenen Ebenen bewegen. Beispiele können Marktdynamiken, Technologiesprünge, Veränderungen der gesellschaftlichen Bewertung einer Technologie oder auch neue Anwendungsformen einer Technologie sein. Die konkrete Form der Szenarien wird in der Literatur, ähnlich wie die Wahl der Methoden zur Technologiefeldanalyse, nicht eingeschränkt. Auf einem Spektrum von narrativen Texten bis hin zu Planspielen sind somit verschiedene Ausprägungen der Szenarien möglich. Zentral für die Entwicklung der Szenarien ist der mit ihnen verfolgte Zweck, der späteren Aktivierung der im Feld aktiven und der sozialen Akteure außerhalb des Technologiefelds. Die soziotechnischen Szenarien sollen für sie Ausgangs- und Bezugspunkt der Reflexion ihrer Strategien sein, und ihnen die damit verbundenen Implikationen für übergreifende, gesellschaftliche Dynamiken und Zusammenhänge aufzeigen. Somit kann es in der Phase der Technologiefeldanalyse ebenfalls von Bedeutung sein, die Realitäten und Alltage der Akteure im Technologiefeld zu erfassen und nachzuvollziehen, um sie bei der Entwicklung der sozio-technischen Szenarien zu berücksichtigen und bewusst abzuwandeln (siehe Abb. 26).

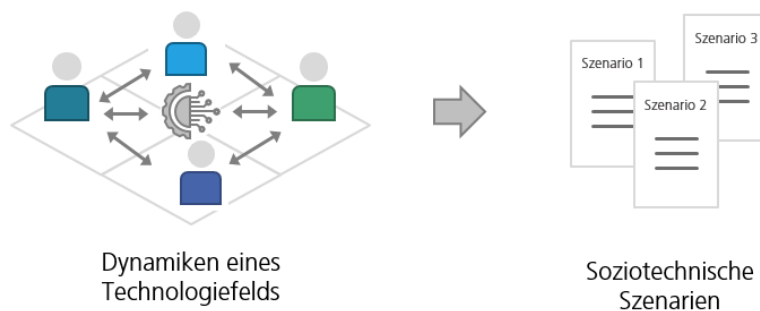


Abbildung 26: Die Erkenntnisse der Technologiefeldanalyse werden in soziotechnische Szenarien überführt

## Durchführung von Workshops

Für die Akteure des Technologiefelds sowie die sozialen Akteure stellen die soziotechnischen Szenarien den Ausgangs- und Bezugspunkt für den gemeinsamen Dialog dar. Der gemeinsame Dialog zur Artikulation von Standpunkten und Strategien soll im Rahmen von Workshops stattfinden, den die CTA agents im Vorfeld vorbereiten und moderierend begleiten. Anhand der Szenarien soll den Akteuren ermöglicht werden, mögliche Entwicklungen des Technologiefelds an fiktiven, aber dennoch denkbaren Beispielen zu reflektieren und mit Aspekten in Berührung zu kommen, die ihre üblichen Gedankengänge bewusst überschreiten. Um möglichst zielführende Diskussionen im Rahmen der Workshops zu ermöglichen, ist für die Organisation der Workshops von hoher Priorität, ein Teilnehmerfeld bestehend aus beiden Akteursgruppen vorzusehen, möglichst im selben Verhältnis von gestaltenden und sozialen Akteuren. Aufschlussreiche Diskussionen verspricht dabei besonders die Rekrutierung sozialer Akteure, die mit Blick auf potenzielle negative Folgen der betrachteten Technologie typischerweise eine Minderheit vertreten. Insofern

finden sich in der Literatur verschiedene Hinweise darauf, in den Workshops Akteure zusammenzuführen, die aus möglichst unterschiedlichen Bereichen des Technologiefelds stammen und deshalb in ihrem bisherigen Alltag wenig bis gar keine Berührungspunkte aufweisen. Aufgabe der CTA agents ist es, für die Workshops ein Umfeld zu schaffen, in dem es sämtlichen Akteuren möglichst leichtfällt, die eigenen Standpunkte darzulegen sowie sich auf die Standpunkte und „Welten“ der weiteren teilnehmenden Akteure einzulassen. In einem solchen „safe space“ soll es den Akteuren ermöglicht werden, die möglichen positiven und negativen gesellschaftlichen Folgen ihres Handelns zu erkennen und gleichzeitig einen Raum zu bilden und zu schaffen, der es ermöglicht, erste Ideen und Vorschläge für in der Zukunft notwendige Veränderungen oder Maßnahmen zu entwickeln. Auf diese Weise wird in CTA angestrebt, Veränderungen in zukünftigen Gestaltungsprozessen zu bewirken, da die gestaltenden Akteure durch die Begegnungen und Diskussionen in den Workshops ein besseres – da ganzheitlicheres – Verständnis über das Ausmaß ihrer Handlungen und Entscheidungen besitzen (siehe Abb. 27).

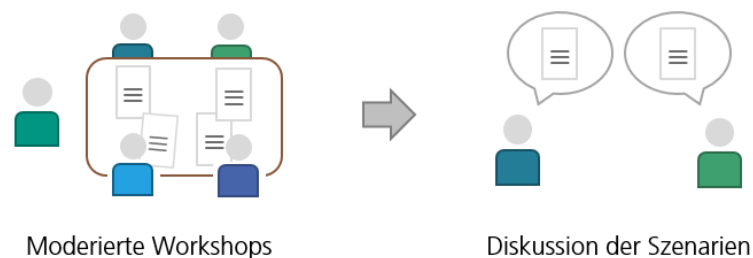


Abbildung 27: In moderierten Workshops diskutieren Akteure die soziotechnischen Szenarien

### 4.3 Value Sensitive Design

Das VSD ist ein Gestaltungsansatz in dem danach gestrebt wird, sämtliche an der Gestaltung von IT-Systemen beteiligten Professionen zur Adressierung auftretender ethischer Fragestellungen zu befähigen. In VSD wird sich Gestaltungsprozessen mit einer integrativen Denkweise angenähert, die zusätzlich zu technischen und sozialen Aspekten besonders die Bedeutungen und Auswirkungen menschlicher Werte (human values) in und auf Gestaltungsprozesse zu ergründen sucht. Durch die Zusammenführung und Weiterentwicklung von Konzepten verschiedener interdisziplinärer Communities strebt VSD nach einer vorausschauenden, werteorientierten Gestaltung, die Fragen mit ethischem Gehalt als bedeutsamen Bestandteil von Gestaltungsprozessen fasst und die Auseinandersetzung mit solchen Fragen systematisch in das Gestaltungsgeschehen zu integrieren sucht:

*„it provides theory, method, and practice to account for human values in a principled and systematic manner throughout the technical design process“*  
(B. Friedman & Hendry, 2019, S. 18).



Zusätzlich sind in VSD verschiedene Überschneidungen zu (C)TA-Arbeit vorhanden, da auch in VSD vorrangig danach gestrebt wird, Reflexionsräume zu schaffen, um den potenziell negativen Folgen von Entscheidungen und Handlungen in Gestaltungsprozessen vorzubeugen.

Die Entwicklung von VSD nimmt ihren Anfang in den frühen 1990er-Jahren in den USA. Zu dieser Zeit beginnt sich der flächendeckende Einsatz von IT-Systemen in professionellen wie auch privaten Lebensbereichen in zunehmendem Maß auszudehnen. Dies trägt dazu bei, dass sich in verschiedenen wissenschaftlichen Communities ein Diskurs zu entfalten beginnt, der sich verstärkt mit den ethischen Implikationen von IT-Systemen auseinandersetzt. An diesem Diskurs beteiligen sich auch Batya Friedman und Peter Kahn, die beide federführend in der Entwicklung von VSD sind und in 2024 beide als Professoren an der University of Washington tätig sind. Friedman und Kahn bringen einen ersten Beitrag in die Debatte ein (B. Friedman & Kahn, 1992), in dem sie die Auswirkungen von Gestaltungsprozessen auf die menschliche Handlungsfähigkeit betrachten. Darin unterstützen sie die aufkommende Bewegung, sich in Feldern wie der HCI mit den ethischen Implikationen der IT-Gestaltung zu befassen:

*„philosophy seldom tells us how or why problems relevant to a philosophical position involving computing occur in practice, let alone what can most effectively resolve them“* (B. Friedman & Kahn, 1992, S. 12).

Vier Jahre später stellt Friedman einen eigenen Ansatz vor, der erstmals den Namen VSD trägt. Der Ansatz zeigt weitere Argumente für die Auseinandersetzung mit Werten auf, um die Gestaltung von IT-Systemen zu fördern, die *„from an ethical position – we can and want to live with“* (B. Friedman, 1996, S. 18). Friedman und ihre Mitstreiter setzen die Erarbeitung der VSD-Theorie fort (B. Friedman, 1997) und stellen sie in verschiedenen Workshops mit interdisziplinärer Beteiligung zur Diskussion (B. Friedman, 1999). Wenig später kann schließlich der Grundstein für eine VSD-Community gelegt werden, in der Expertisen aus Fachbereichen wie den Computer Ethics, des Interaction Designs, der Softwareentwicklung, der Philosophie und des Participatory Designs (PD) vertreten sind. Noch im alten Jahrtausend gründen Friedman und Kahn unter Mitwirkung von Alan Borning an der University of Washington das Value Sensitive Design Lab, das dort an der Schnittstelle der Information School, dem Department of Psychology und dem Department of Computer Science and Engineering angesiedelt ist, und den drei Wissenschaftlern als Arbeitsstätte zur Weiterentwicklung des VSD-Ansatzes dient (B. Friedman et al., 2002). Ein erster wichtiger Meilenstein der Weiterentwicklung stellt ein Kapitel im Human Computer Interaction Handbook dar (B. Friedman & Kahn, 2007), in dem VSD erstmals auch an den erweiterten Diskurs zur wertorientierten Gestaltung technischer Systeme anknüpft. Dies erfolgt, indem die Autoren in ihrem Beitrag VSD in Beziehung zu Aspekten setzen, die typischerweise in Debatten zur Berufsethik diskutiert werden. VSD findet zudem Eingang in die Enzyklopädie der HCI (B. Friedman, 2004) und in 2006 stehen erste konkretere Erkenntnisse zu VSD-Heuristiken zur Verfügung, die in Fallstudien erprobt wurden (B. Friedman, Kahn, Hagman, et al., 2006). Auch beginnen Friedman, Kahn und Borning damit, die Ziele und Inhalte von VSD erstmals an nachfolgende Generationen zu vermit-

teln (Davis & Nathan, 2015). Ähnlich wie CTA erfährt auch VSD zu diesem Zeitpunkt zunehmende internationale Aufmerksamkeit, die ihm aufgrund seiner ganzheitlichen Betrachtung der werteorientierten Gestaltung entgegengebracht wird (Brey, 2010; Manders-Huits, 2011). In Europa tragen besonders Wissenschaftler der Technischen Universität Delft zur Weiterentwicklung von VSD bei (van den Hoven, 2013; van Wynsberghe & Robbins, 2014). So gewinnt VSD auch außerhalb informatischer Communities zunehmend an Resonanz. Im Jahr 2024 besitzt VSD nach wie vor eine hohe Relevanz im interdisziplinären Diskurs werteorientierter Gestaltung (z.B. Korenhof, 2023; Umbrello & Gambelin, 2023; Zuber et al., 2024). Auch hinsichtlich der (menschzentrierten) Gestaltung von KI-Systemen finden sich verschiedene Studien, die über Erkenntnisse der Anwendung von VSD berichten (z.B. Sadek et al., 2023; Solanki et al., 2023).

#### 4.3.1 Zentrale Annahmen und Konzepte

Die Theorie auf die sich VSD gründet besteht aus einer intellectual landscape der vier Communities (1) applied moral philosophy, (2) social informatics, (3) CSCW und (4) PD (B. Friedman & Hendry, 2019, S. 23–24). Aus diesen Communities werden in VSD verschiedene Erkenntnisse und Schwerpunkte aufgegriffen und zum Zweck der werteorientierten Gestaltung neu zusammengefügt. Während die letzten drei genannten Communities VSD insbesondere zu Aspekten der Gestaltung und der Aneignung von IT-Systemen als Informationsquelle dienen, informiert die applied moral philosophy in VSD wesentlich beim Anstellen philosophischer Betrachtungen. Besonders die Formulierung einer Arbeitsdefinition zu menschlichen Werten steht dabei im Vordergrund, die inzwischen definiert sind als „what is important to people in their lives, with a focus on ethics and morality“ (B. Friedman & Hendry, 2019, S. 31). Hieran wird die pragmatische Denkweise in VSD erkennbar, die neben einer übergreifenden Auseinandersetzung auch die Auseinandersetzung mit spezifischen Werten ermöglicht, die sich in Gestaltungskontexten als besonders bedeutsam erweisen können. Als Beispiele für derartige Werte nennt die Literatur freedom from bias, privacy, human autonomy oder auch human welfare. Auch wenn in VSD ein pragmatischer Zugang zu menschlichen Werten vorherrscht, wird die für die Auseinandersetzung mit Werten notwendige Reflexionsfähigkeit nicht vernachlässigt. Besonders deutlich wird dies an einer Erläuterung Batya Friedmans, aus der hervorgeht, weshalb in VSD das Konzept der human values den Vorzug vor dem der moral values erhält:

*„to highlight the complexity of social life, and to provide a basis for analyses wherein personal and conventional values can become morally implicated“*  
(B. Friedman, 1997, S. 5).

Menschliche Werte werden in VSD nicht als abstrakte, losgelöste Konstrukte betrachtet, sondern als sozial und kulturell eingebettete Gebilde, die aufeinander wechselwirken:

*„human values do not exist in isolation. Rather, much like the threads in a spider web, values are situated in a delicate balance“* (B. Friedman & Hendry, 2019, S. 44)

Eine entsprechend hohe Aufmerksamkeit wird in VSD deshalb der Erfassung und Ergründung von Wertekonflikten (value tensions) gewidmet. Diese werden als Gestaltungsprozesse durchgängig begleitende Phänomene betrachtet, die es gilt (1) grundsätzlich anzuerkennen, (2) unter den konfligierenden Werten die relevantesten zu identifizieren und (3) idealerweise unter ihnen durch die Anwendung der in VSD vorgeschlagenen Methoden ein Gleichgewicht zu erzielen. An vielen Stellen der VSD-Literatur wird darauf hingewiesen, dass es sich bei den dazu notwendigen Aktivitäten um herausfordernde Aktivitäten für sämtliche am Gestaltungsprozess beteiligten Personen handelt. Jeder Gestaltungsprozess weist einzigartige Merkmale und Dynamiken auf, die Einfluss auf die Interpretationen und Bedeutungen von Werten zur Folge haben kann. Die Erfassung der Spezifika einer solchen Situiertheit von Gestaltungsprozessen ist eine der zentralen Zielsetzungen von VSD, auf die besonders die Auslegung des mittleren Namensbestandteils „sensitive“ hinweisen soll. Ein weiteres grundlegendes Element der VSD-Theorie stellt die interactional stance dar, welche die in VSD vertretene Perspektive auf Mensch-Technik-Verhältnisse sowie die Aneignung von Technik durch Menschen wiedergibt. In VSD wird der Mensch in einer interaktiven Dualität begriffen, einer Art kontinuierlichem Kreislauf, da er zur Bewältigung seiner Lebensumstände einerseits Werkzeuge und technische Artefakte gestaltet, die seine Lebensumstände verändern und die wiederum die Gestaltung neuer Werkzeuge zur Folge hat (B. Friedman & Hendry, 2019, S. 35). Durch die Anerkennung dieses wechselwirkenden Verhältnisses von Mensch und Technik wird sich in VSD (wie auch in CTA) deutlich von technikdeterministischen Ansätzen abgegrenzt und das Ignorieren sozialer Aneignungsprozesse während der Gestaltung wird als problematisch betrachtet (Davis & Nathan, 2015, S. 15). Nach Ansicht in VSD soll der design space, der gedankliche und physische Raum, der bei der Gestaltung technischer Systeme zur Verfügung steht, als soziotechnischer Raum begriffen werden, besonders um Gestaltern Zugang zu Perspektiven und Inspirationen zu ermöglichen, die ihnen bei einer ausschließlich technischen Betrachtung verwehrt bleiben würden. Daraus leitet sich auch ein Anknüpfungspunkt zu Bestrebungen der TA ab, denn ähnlich wie in TA wird auch in VSD die Verantwortung für nicht-intendierte Folgen bei den für die Gestaltung verantwortlichen Personen verortet:

*„the interactional stance implies that the design team needs to be alert and ready for unintended consequences, which may arise narrowly with specific tools or individual practices or more broadly with infrastructure and social practice“* (B. Friedman & Hendry, 2019, S. 36)

Der in VSD durchgehend referenzierte Standpunkt „progress, not perfection“ ist ein weiteres charakteristischstes Merkmal in VSD, da es erneut den pragmatischen Charakter des Ansatzes unterstreicht und sehr deutlich wiedergibt. Ausgangspunkt dieses Standpunkts ist die Anerkennung der hohen inhaltlichen Komplexität, die in vielen Gestaltungsprozessen vorherrscht und mit der sich entsprechend viele Gestalter konfrontiert sehen. Sie haben einerseits zur Aufgabe, mit limitiert zur Verfügung stehenden Ressourcen technische Systeme zu gestalten, und zusätzlich dem Anspruch gerecht zu werden, die ethischen Implikationen ihrer Aktivitäten und Systeme abzuschätzen. In VSD wird die Ansicht vertreten, dass Gestalter, die unter solchen Bedingungen tätig sind, Instrumente und Orientierungshilfen zur Verfügung gestellt werden sollten, die es ihnen erlauben, *trotz* offener Fragen zu ethischen Sachverhalten auf gestalterischer Ebene Fortschritte zu erzielen (progress). Diese Funktion wird in VSD zu erfüllen versucht, indem Gestalter durch die vorgeschlagenen Methoden ermöglicht wird, Defizite von Gestaltungsprozessen auszumachen und Lösungswege zu deren Adressierung aufzuzeigen. Somit wird auch in VSD danach gestrebt, Gestalter zu einem Wortschatz zu verhelfen, der es ihnen ermöglichen soll, die oftmals weitreichenden sozialen Folgen ihrer Arbeit mit der Öffentlichkeit zu diskutieren. In dieser Befähigung zur Durchführung eines öffentlichen Diskurses liegt somit ein weiterer Überschneidungspunkt von VSD- und CTA-Interessen. Durch das Konstrukt des progress not perfection werden in VSD die Rahmenbedingungen anerkannt, denen viele Gestaltungsprozesse in der Realität unterliegen und es wird versucht, in diesen bereits komplexen Gemengelagen Räume zu schaffen, die eine (zusätzliche) Auseinandersetzung mit ethischen Aspekten ermöglichen. Die Befürworter von VSD lassen sich in dieser Bestrebung nicht von der in ethischen Diskursen oftmals angebrachten Kritik der Unvollständigkeit zurückhalten, da in VSD eine restlos befriedigende Klärung ethischer Fragen (perfection) nur schwerlich als zu erwirken erachtet wird. Aus diesem Grund wird in VSD die Überzeugung vertreten, dass Gestaltungsansätze, die sich der Auseinandersetzung mit ethischen Aspekten widmen, „need not require perfection, but commitment to practice. And through practice, progress“ (B. Friedman & Hendry, 2019, S. 51).

Die interactional stance in Verbindung mit der Zielsetzung der systematischen Berücksichtigung menschlicher Werte erfordert es in VSD darzulegen, anhand welcher Kriterien die Bedeutsamkeit von Werten ermittelt und gewichtet werden kann. Die VSD-Position hierzu lautet, dass dies erfolgen kann, indem im Gestaltungsprozess die Menschen fokussiert werden, „who are or will be significantly implicated by the technology“ (B. Friedman & Hendry, 2019, S. 39). Die damit verbundene Annahme lautet, dass die Werte dieser Menschen aufgrund ihrer Betroffenheit von den Auswirkungen eines technischen Systems eine hohe Bedeutung im Gestaltungsprozess erfahren sollten. Dieser Standpunkt berührt jedoch eine sehr hohe inhaltliche Komplexität, denn er kann beispielsweise eine Argumentation dafür erforderlich machen, weshalb bestimmte Gruppen (und somit Werte) in einem Gestaltungsprozess als nicht bedeutsam erachtet werden:

*„How do you define “implicated?” And what makes an implication significant enough to warrant consideration in the design process?“* (B. Friedman & Hendry, 2019, S. 39).

Ferner handelt es sich hierbei um die Herausforderung, die (Verantwortungs-)Grenzen eines Gestaltungsprozesses abzuwägen und zu definieren, sowie die Identifikation von signifikant betroffenen Menschen auf möglichst transparente und nachvollziehbare Weise durchzuführen. Mit der Definition von Grenzen ist unweigerlich das Eingehen von Kompromissen verbunden – ein Umstand der in VSD durchaus anerkannt wird, dem jedoch mit Verweis auf das progress, not perfection zu begegnen gesucht wird:

*„While such questions defy definitive responses, further theoretical constructs in value sensitive design point the way forward“ (B. Friedman & Hendry, 2019, S. 39).*

In VSD wird deshalb als ein weiteres zentrales theoretisches Element das Rollenkonzept der Stakeholder verwendet, mit dessen Hilfe zwischen von Technik direkt betroffenen (direct stakeholders) und indirekt betroffenen (indirect stakeholders) Stakeholdern unterschieden wird, um die verschiedenen bedeutsamen Zielgruppen und ihre Werte zu ermitteln und zu beschreiben. An diesem Element wird eine weitere Parallele zu TA erkennbar, denn wie auch in vielen Kontexten von TA-Arbeit erfahren in VSD mit den indirekten Stakeholdern besonderes diejenigen Gruppen eine hohe Bedeutung, die im Gestaltungsgeschehen aufgrund dem Fehlen offensichtlicher Ursache-Wirkungszusammenhänge typischerweise wenig bedacht werden.

### 4.3.2 Vorgeschlagene Methodik

Die zentralen Annahmen und Konzepte werden in VSD auf unterschiedliche Weise auf die Methodik übertragen. Die tripartite methodology bezeichnet das grundsätzliche methodische Framework von VSD, das grundlegend zwischen den drei Untersuchungstypen der conceptual, empirical and technical investigations unterscheidet (siehe Abb. 28).

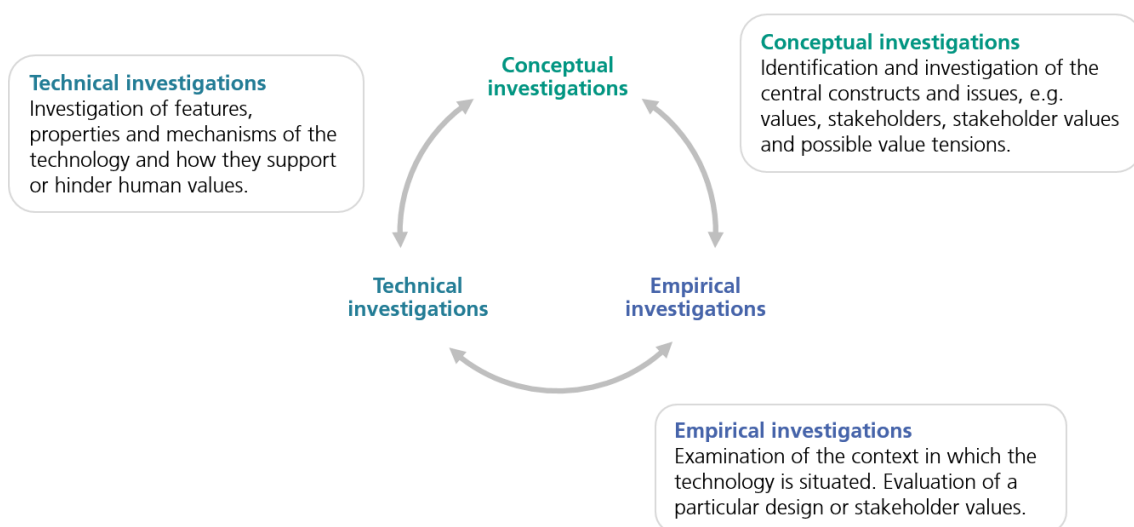


Abbildung 28: Die tripartite methodology in VSD

Diese drei Untersuchungstypen verhalten sich im Verständnis von VSD komplementär zueinander und sollen ermöglichen, durch die Anwendung unterschiedlicher methodischer Mittel die unterschiedlichen Ebenen der Wertefragen betreffenden Aspekte zu beleuchten:

*„No one type of investigation is sufficient on its own; rather, all three investigation types are needed to inform and shape and reshape each other“* (B. Friedman & Hendry, 2019, S. 39)

### Conceptual investigations

Die conceptual investigations dienen der Erhebung der mehr theoretischen und (Gestaltungs-)philosophischen Aspekte, die in Gestaltungsprozessen zum Tragen kommen. Typischerweise dienen sie dazu, (1) die in bzw. für einen Gestaltungsprozess zentralen menschlichen Werte zu ermitteln und zu beschreiben, sowie (2) die direkten und indirekten Stakeholder zu identifizieren. Auch wenn der Zeitpunkt für die Durchführung von conceptual investigations frei gewählt werden kann, empfiehlt es sich, conceptual investigations zu Beginn eines Gestaltungsprojekts bzw. einer VSD-Untersuchung durchzuführen, da sie aufgrund der Ermittlung konzeptueller Sachverhalte Erkenntnisse hervorbringen, die durchgängig charakteristisch für einen Gestaltungsprozess sind.

Zur Ermittlung zentraler Werte werden in VSD üblicherweise Fragen offenen Typs gestellt, wie beispielsweise: „What conceptual models, if any, for operationalizing a given value or values will be employed?“ (B. Friedman & Hendry, 2019, S. 37). Die Beantwortung derartiger Fragen sollen die tieferliegenden Zusammenhänge und Motivationen der involvierten Gruppen offenlegen. Solche Leitfragen ermöglichen es, zielgerichtet zentrale Werte zu identifizieren und nach deren Identifikation, durch die Formulierung von Arbeitsdefinitionen für besagte Werte, zu ihrer Explikation im Gestaltungsprozess beizutragen. In der Formulierung gestaltungsprozessspezifischer Arbeitsdefinitionen liegt für VSD-Projekte ein besonderer Wert, da sie sich aufgrund gegebener Spezifika an entscheidenden Stellen von etablierten Wertdefinitionen unterscheiden können, oder es erlauben, sich auf bestimmte Teilaspekte zu fokussieren. Durch die Erarbeitung von Arbeitsdefinitionen soll zudem ermöglicht werden, Wertekonflikte (frühzeitig) sichtbar zu machen.

Um die direkten und indirekten Stakeholder eines Gestaltungsprozesses zu ermitteln, ermöglichen Methoden der conceptual investigations eine gezielte Annäherung an den (vermeintlichen) Nutzungskontext. Dies erfolgt durch die Auseinandersetzung mit sämtlichen Stakeholdergruppen im Feld und außerhalb des Felds und ermöglicht so Stakeholder auszumachen, die regelmäßig und auf bedeutsame Weise in Interaktion mit einem System treten. Auch können dadurch diejenigen Stakeholder identifiziert werden, deren Lebensumstände sich möglicherweise zufällig oder unerwartet durch die Einführung, Nutzung oder Adaption eines Systems verändern.

## Empirical investigations

Die empirical investigations verfolgen das Ziel, durch die Anwendung quantitativer und qualitativer Methoden empirisches Datenmaterial über den Nutzungskontext zu erheben, um detailliertere Informationen über die dort aktiven Stakeholder und ihre Werte zu ermitteln. Sie fokussieren sich darauf, die Erfahrungen der Stakeholder im Umgang mit Technik sowie ihre Werteinterpretationen zu untersuchen und schärfen dadurch üblicherweise die Erkenntnisse der conceptual investigations. Ähnlich wie die conceptual investigations werden auch die empirical investigations in der Regel von übergeordneten Fragen geleitet, allerdings zielt die Art ihrer Fragestellungen auf vergleichsweise konkretere Aspekte ab, z.B.: „Are there differences between espoused practice (what people say) compared with actual practice (what people do)?“ (B. Friedman & Hendry, 2019, S. 38). Auch für die Durchführung von empirical investigations werden in VSD verschiedene Methoden zur Verfügung gestellt – Ziel der empirical investigations soll sein, relevante Handlungen, Werteverständnisse und Bedenken der Stakeholder möglichst authentisch, und möglichst im Nutzungskontext zu erfassen. So werden in VSD-Studien beispielsweise Fragebögen, Interviews, Beobachtungen oder Artefaktanalysen durchgeführt.

## Technical investigations

Bei den technical investigations steht die konkrete Technik, also das technische Artefakt oder System im Vordergrund. Technical investigations werden von der Zielsetzung geleitet, zu ergründen, inwieweit bestimmte technische Eigenschaften, Komponenten oder Funktionen bestimmte menschliche Werte fördern und unterstützen, und inwieweit sie Konfliktpotenzial für bestimmte Werte besitzen:

*„What features of a technical infrastructure enable, hinder, or even foreclose certain kinds of designs for supporting human activity?“* (B. Friedman & Hendry, 2019, S. 38)

Dazu verweist die VSD-Literatur auf den Erfahrungswert, dass eine Kombination aus zeitlich vorwärts- und rückwärtsgewandten Untersuchungen aufschlussreich sein kann, beispielsweise um Kenntnisse über historische Ursachen für bestimmte Systemeigenschaften zu erhalten, die Aufschluss über gegenwärtige Werteinterpretationen von Stakeholdern geben. Die technical investigations können jedoch auch dazu genutzt werden, proaktive Gestaltungsmaßnahmen zu identifizieren und vorzubereiten, beispielsweise um als bedeutsam identifizierte Werte technisch künftig stärker zu unterstützen oder um vorhandene Wertekonflikte auf technischer Ebene zu adressieren.

Durch die Verbindung konzeptueller, empirischer und technischer Betrachtungen soll in VSD-Projekten die Erarbeitung eines möglichst ganzheitlichen Bildes eines Gestaltungsprozesses erfolgen, die anschließend eine informierte Entscheidungsfindung bezüglich seiner Wertefragen ermöglicht. Hinsichtlich Dauer, Intensität, Zeitpunkt, Anzahl oder Reihenfolge der Untersuchungen werden in der VSD-Literatur keine Einschränkungen vorgenommen, um Anwendern bei der Auswahl und Durchführung der Methoden ausreichend Raum zu

geben, den spezifischen Gegebenheiten ihres jeweiligen Gestaltungsprozesses Rechnung zu tragen. Auch wird in VSD nicht der Anspruch erhoben, ausschließlich die in der VSD-Literatur referenzierten Methoden für die verschiedenen Untersuchungen zu verwenden. Stattdessen ist VSD auf eine Weise konzipiert worden, um „in concert with, rather than to supplant, other good technical practices“ (B. Friedman & Hendry, 2019, S. 85) verwendet zu werden. Auch wenn VSD initial für die Gestaltung von IT-Systemen konzipiert wurde, so ermöglicht es die tripartite methodology dennoch, dass VSD auch auf die Gestaltung nicht-informatischer Systeme angewendet werden kann. Ähnlich wie bei CTA sind den Anwendungsbereichen von VSD somit nur wenig Grenzen gesetzt.

## Pool von Gestaltungsmethoden

Mitte 2024 verweist das aktuellste Standardwerk zu VSD (B. Friedman & Hendry, 2019) auf insgesamt 17 Methoden, die in VSD entwickelt bzw. für VSD-Zwecke adaptiert wurden (siehe Abb. 29).

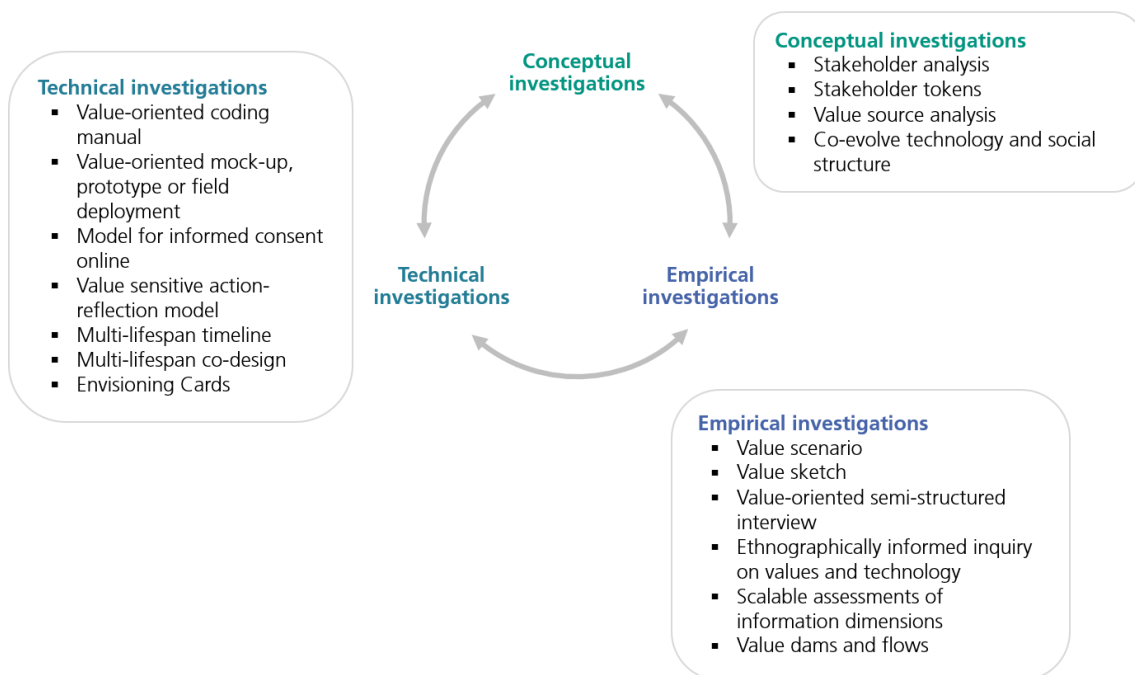


Abbildung 29: Zuordnung der VSD-Methoden zu den drei Untersuchungstypen der tripartite methodology

Diese 17 Methoden werden im genannten Standardwerk dargelegt, sodass ich für die Dissertation darauf verzichte, jede einzelne der Methoden im Detail zu beschreiben. Jedoch erfordern die drei Methoden (1) der Stakeholder analysis (conceptual investigations), (2) der Value dams and flows (empirical investigations) und (3) des Value-oriented prototyping (technical investigations) eine genauere Betrachtung, was durch das Studiendesign der Expertenevaluation von CTA und VSD bedingt ist (siehe Kap. 6.2). Nach ausführlicher



Auseinandersetzung mit den verschiedenen VSD-Methoden schienen mir diese drei Methoden für die Expertenevaluation geeignet zu sein, da sie je einen Untersuchungstypus der tripartite methodology repräsentieren.

Eine Stakeholder analysis (conceptual investigation) hat zum Ziel, direkte und indirekte Stakeholder, nicht-intendierte Folgen sowie potenzielle Nutzen und Risiken eines technischen Systems zu identifizieren (B. Friedman, Kahn, Hagman, et al., 2006; Nathan et al., 2008). Durchgeführt wird die Methode in der Regel von Personen aus dem Projektteam, das für die Gestaltung eines Systems verantwortlich ist. Das Projektteam erstellt für die Stakeholderanalyse zwei Listen, eine Liste zu direkten Stakeholdern und eine Liste zu indirekten Stakeholdern. Direkte Stakeholder beschreiben Personen oder Gruppen, die direkt mit dem System interagieren werden. Indirekte Stakeholder beschreiben Personen oder Gruppen, die nicht direkt mit dem System interagieren werden, die in ihrem Leben jedoch Auswirkungen erfahren werden, die durch die direkten Stakeholder und ihren Interaktionen mit dem System hervorgerufen werden (siehe Abb. 30).

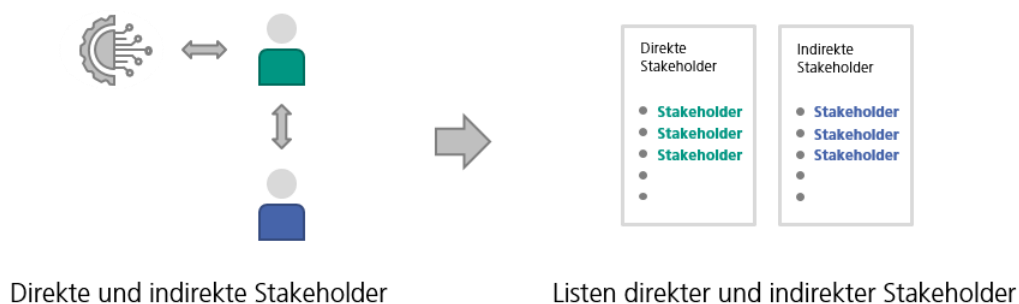
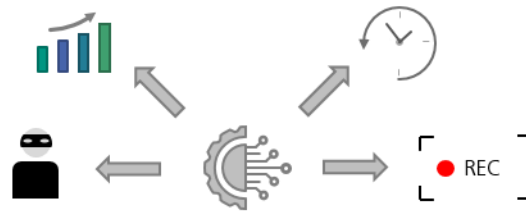


Abbildung 30: Identifikation direkter und indirekter Stakeholder  
(Stakeholderanalyse)

Das Projektteam identifiziert daraufhin nicht-intendierte Nutzungskontexte und Folgen des Systems. Diese beschreiben Nutzungskontexte, in denen die Nutzung des Systems vorstellbar ist, durch die sowohl positive wie auch negative Effekte hervorgerufen werden können und denen im bisherigen Gestaltungsprozess nur wenig Beachtung geschenkt wurde (siehe Abb. 31).



### Nicht-intendierte Nutzungskontexte und Effekte eines Systems (Beispiele)

Abbildung 31: Identifikation möglicher nicht-intendierter Nutzungskontexte und Folgen (Stakeholderanalyse)

Das Projektteam identifiziert ferner potenzielle Nutzen und Risiken des Systems. Nutzen beschreiben dabei positive Effekte oder Veränderungen für Stakeholder und Risiken mögliche Schäden oder Nachteile für Stakeholder, die durch das System entstehen können (siehe Abb. 32).

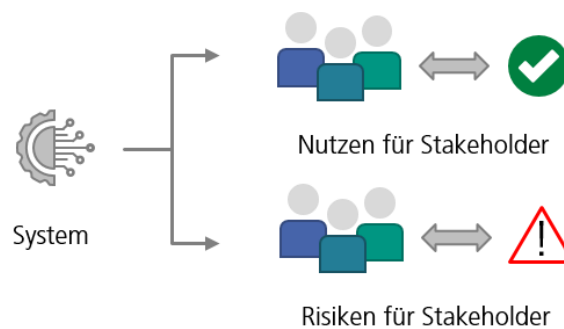


Abbildung 32: Identifikation potenzieller Nutzen und Risiken für Stakeholder (Stakeholderanalyse)

Bei der Methode Value dams and flows (empirical investigation) wird mithilfe eines Fragebogens ermittelt, wie Stakeholder mögliche Nutzen und Risiken eines Systems beurteilen (Denning et al., 2010; Miller et al., 2007). Durchgeführt wird die Methode in der Regel von Personen aus dem Projektteam, das für die Gestaltung eines Systems verantwortlich ist. Zu Beginn identifiziert das Projektteam in einer internen Analyse potenzielle Nutzen und Risiken des Systems (analog zur Stakeholderanalyse, siehe Abb. 32). Nutzen beschreiben dabei positive Effekte oder Veränderungen für Stakeholder, während Risiken mögliche Schäden oder Nachteile beschreiben, die Stakeholdern durch das System entstehen können. Das Projektteam erstellt anschließend einen Fragebogen, der Aussagen über die zuvor identifizierten Nutzen und Risiken enthält. Jeder Aussage sind durch eine Likert-Skala mehrere Beurteilungsmöglichkeiten zugeordnet, die von „stimme überhaupt nicht zu“ bis „stimme voll und ganz zu“ reichen (siehe Abb. 33). Die Stakeholder werden anschließend gebeten, den Fragebogen zu bearbeiten und jede Aussage zu bewerten.



Abbildung 33: Befragung von Stakeholdern zu potenziellen Nutzen und Risiken (Value dams and flows)

Das Projektteam wertet die Fragebögen anschließend aus und ermittelt anhand der Ergebnisse die namensgebenden Value dams und Value flows. Value dams bezeichnen dabei Aussagen zu potenziellen Risiken des Systems, die mindestens 10% der Stakeholder mit „**stimme voll und ganz zu**“ beurteilen. Value flows bezeichnen Aussagen zu potenziellen Nutzen des Systems, die mindestens 50% der Stakeholder entweder mit „**stimme zu**“ oder „**stimme voll und ganz zu**“ beurteilen (siehe Abb. 34).

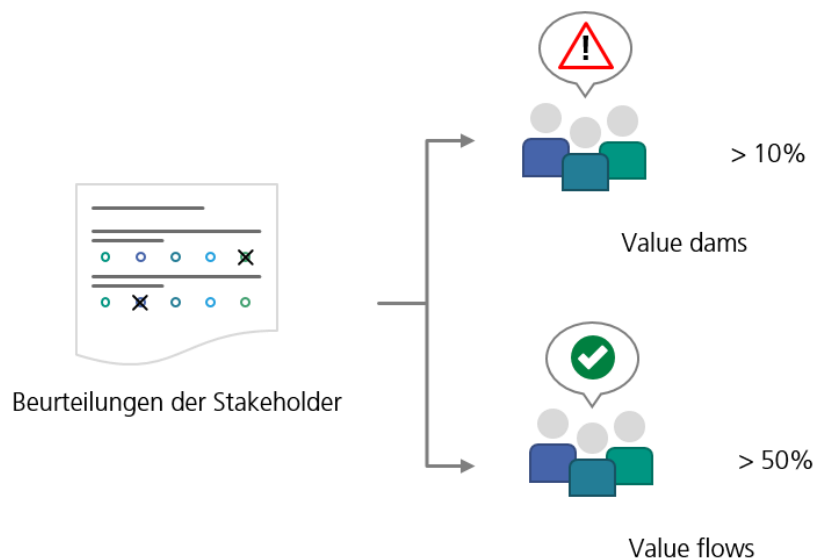


Abbildung 34: Ermittlung von Value dams und Value flows (Value dams and flows)

Beim Value-oriented Prototyping (technical investigation) erstellen Stakeholder Prototypen des Systems, die ihre Wertvorstellungen repräsentieren und ihr Verständnis der Funktionsweise des Systems veranschaulichen (Woelfer & Hendry, 2009; Yoo et al., 2013). Durchgeführt wird die Methode in der Regel von Personen aus dem Projektteam, das für die Gestaltung eines Systems verantwortlich ist. In einem Workshop sollen Stakeholder Prototypen entwerfen und realisieren, die ihre Vorstellungen von der Funktionsweise des Systems (oder einzelner Systemkomponenten) repräsentieren.

Zur Erstellung der Prototypen stellt das Projektteam verschiedene Gegenstände und Bastelmaterialien zur Verfügung. Anschließend an das Prototyping erläutern und demonstrieren die Stakeholder die Funktionsweise ihrer Prototypen (siehe Abb. 35).

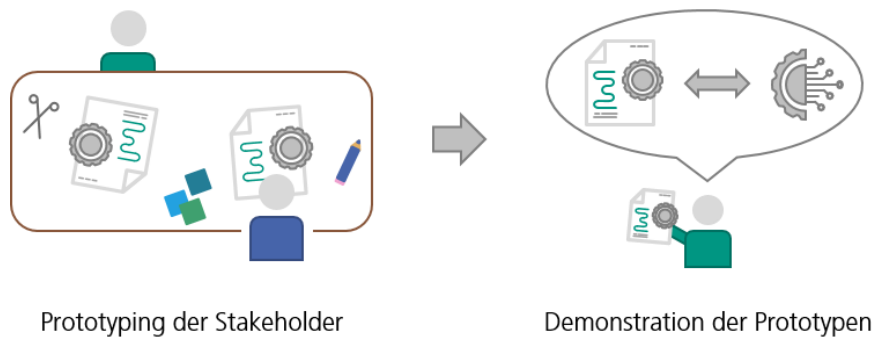


Abbildung 35: Durchführung eines Prototyping-Workshops (Value-oriented Prototyping)

Das Projektteam hält diese Ausführungen und Demonstrationen fest und analysiert die Prototypen anschließend dahingehend, welche Werte den Stakeholdern mit Blick auf das System wichtig sind und inwieweit das System im aktuellen Zustand diese Werte unterstützt oder beeinträchtigt. Bei vorhandenen Wertekonflikten zwischen den Werten der Stakeholder und dem Projektteam findet das Projektteam Wege, mit den Wertekonflikten umzugehen (siehe Abb. 36).



Abbildung 36: Analyse der Prototypen (Value-oriented Prototyping)

## 4.4 Zwischenbetrachtung

Nach der grundlegenden Auseinandersetzung mit den beiden Ansätzen CTA und VSD führe ich an dieser Stelle eine Zwischenbetrachtung durch, in der ich (vorerst noch auf recht abstrakter Ebene) verschiedene Aspekte diskutiere, die für die Fragestellung der Dissertation bedeutsam sind. Diese Zwischenbetrachtung erfolgt auch vor dem Hintergrund der Tatsache, dass VSD in 2024 über die angestammte Community hinaus große Reputation genießt und in verschiedenen technikphilosophischen Communities Rezeption erfährt, die sich mit ethischen Fragestellungen und Implikationen von Technikgestaltungsprozessen auseinandersetzen. CTA hingegen wird überwiegend in TA-Arbeiten referenziert und hat außerhalb der TA-Community kaum an Bedeutung hinzugewonnen.

### 4.4.1 Abstraktionsebene und Blick auf Gestaltungsprozesse

Was beim Vergleich der beiden Ansätze augenscheinlich wird, ist ihre unterschiedliche Abstraktionsebene, ihr unterschiedlicher Blickwinkel auf Gestaltungsprozesse und ihr methodischer Zugang, den sie als Folge dessen wählen und entwickeln. Durch seinen Ursprung in der TA und der sozialwissenschaftlichen Technikforschung werden in CTA Gestaltungsprozesse primär in ihrer gesellschaftlichen Bedeutung betrachtet – maßgebliche Motivation der soft interventions in CTA ist die stärkere Ausrichtung von Gestaltungsprozessen auf übergreifende gesellschaftliche Interessen. Die Notwendigkeit dieser Ausrichtung wird im Kontext der Debatte zu einem „managing technology in society“ (Rip et al., 1995) nachvollziehbar dargelegt. Die Maßnahmen zur Realisierung eines Technologiemanagements für und durch die Gesellschaft sehen jedoch eine bewusst gewahrte Distanz zum tatsächlichen Gestaltungsgeschehen vor, die nur punktuell, und auch dann nur sehr geringfügig zum Zwecke der Durchführung von Workshops verringert wird. In dieser Annäherung zum Gestaltungsgeschehen besteht jedoch bereits der innovative Charakter und das Alleinstellungsmerkmal von CTA für die TA-Community, denn diese besitzt traditionellerweise weniger den Anspruch (und folglich weniger Interesse daran), durch aktives Eingreifen direkte Veränderungen in den Gestaltungsprozessen zu bewirken. In und durch CTA wird diese Einstellung ein Stück weit aufgebrochen und durch für in TA übliche Argumente und Argumentationsweisen dargelegt, dass Gestaltungs- und Entscheidungsprozesse durchaus zum Handlungsfeld von TA-Arbeit gezählt werden können, um dort Aspekte zu Technikfolgen anzubringen und ihre Reflexion und Diskussion zu ermöglichen. In dieser Position steckt somit ein Zugeständnis, dass das in TA ansonsten übliche Selbstverständnis, zu Gestaltungsprozessen mehr indirekt über ihren Auftrag der Politikberatung und den daraus resultierenden Regulierungsmaßnahmen beizutragen, deutlich erweitert, und Technikfolgenabschätzer deutlich (pro-)aktiver als handelnde Akteure positioniert. Somit soll CTA auch ermöglichen, Technikfolgen zeitlich stärker aus einer ex ante-Position abschätzen zu können, da Regulierungsmaßnahmen häufig einen bereits erreichten Status Quo von Technik zu korrigieren suchen. Trotz dieser Zugeständnisse und Bestrebungen werden in der CTA-Methodik mehr indirekte Interventionsmethoden verwendet, die aufgrund ihrer Interventionsintentionen nur in geringem Maße das Potenzial besitzen, direkte Veränderungen in Entscheidungspro-

zessen herbeizuführen. Die Verbindlichkeit der in den CTA-Workshops erarbeiteten Ergebnisse und Commitments, sofern letztere überhaupt von den Akteuren formuliert werden, ist für die Workshopteilnehmer eher gering und erfolgt höchstens auf freiwilliger Basis. Somit wird auch in CTA, trotz der dargelegten Zugeständnisse und Ansprüche, weitestgehend in der typischen TA-Position verblieben – Anwender von CTA sind begleitend und beratend tätig, nicht gestaltend.

Im Unterschied zu CTA ist der Ansatzpunkt in VSD nicht die gesellschaftliche Ebene, sondern die Interaktionsebene zwischen Mensch und Technik, wodurch sich in VSD deutlich näher am aktiven Gestaltungsgeschehen verortet wird. Folglich sind in VSD auch deutlich weniger Vorbehalte vorhanden, direkten Einfluss auf Gestaltungsprozesse zu nehmen. Durch die traditionelle Nähe zur Informatik und zu den Designwissenschaften ist in VSD die Bestrebung zu aktivem Gestaltungshandeln deutlich ausgeprägter, was besonders durch die Position des progress not perfection deutlich wird. Sie repräsentiert das in VSD vorhandene Wissen über den Wunsch von Gestaltern und die ihnen häufig entgegengebrachte Erwartungshaltung, kontinuierlich Fortschritte im Gestaltungsprozess zu erzielen. In VSD wird danach gestrebt, Gestaltern diesem Ziel näher zu bringen, indem ihnen durch verschiedene Methoden Hilfsmittel zur Verfügung gestellt werden, die es ihnen erlauben, sich auf angemessene, aber dennoch pragmatische Weise mit den ethischen Implikationen ihrer Arbeit auseinanderzusetzen. Dadurch wendet sich VSD Gestaltungsprozessen auf einer deutlich niedrigeren Abstraktionsebene zu (siehe Abb. 37).

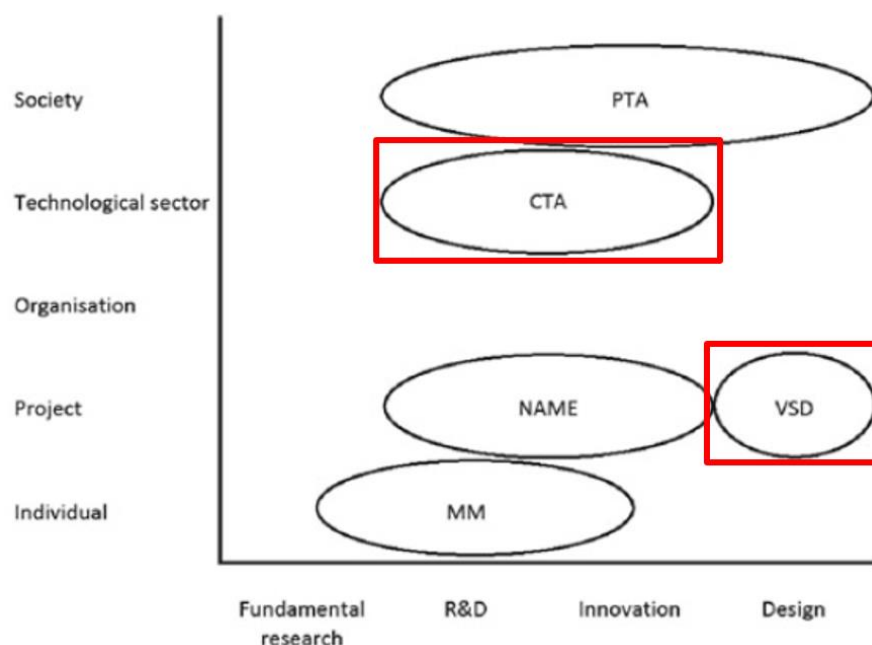


Abbildung 37: In CTA und VSD werden beim Blick auf Gestaltungsprozesse unterschiedliche Abstraktionsebenen verwendet (Doorn et al., 2013, S. 239)

In VSD stellt sich somit weniger die Frage, ob eine Intervention in Gestaltungsprozesse vertretbar ist. Es wird gedanklich direkt einen Schritt weitergegangen und sich den Fragen und Herausforderungen zugewendet, die eine Identifikation und Deckung konkreter Bedarfe von Gestaltern ermöglichen. Dadurch können in VSD deutlich klarer umrissene Vorstellungen von Ziel- und Adressatengruppen entwickelt werden und Sachverhalte als gegeben hingenommen werden, die in CTA aufgrund der dort gewährten Distanz weniger bekannt sein können und vergleichsweise fremd wirken würden. Aufgrund seiner größeren Nähe zu Gestaltungsprozessen sind in VSD auch deutlich weniger Vorbehalte vorhanden, die Verantwortung für die Folgen zu übernehmen, die die Anwendung von Gestaltungsmethoden in und auf Gestaltungsprozesse bewirken. Die VSD-Methoden greifen aktiver in das Gestaltungsgeschehen ein und haben häufig konkrete Entscheidungen zum Ergebnis, wodurch in VSD im Vergleich zu CTA über deutlich mehr Bereitschaft verfügt werden muss, für diese Entscheidungen und besonders die Art und Weise ihrer Herbeiführung Rechenschaft abzulegen. Entsprechend dienen viele Bemühungen der VSD-Methoden dazu, Gestaltungsentscheidungen mithilfe möglichst vieler Informationen auf möglichst transparente Weise zu fällen, und bei den damit verbundenen Kompromissfindungen möglichst viele Beteiligte und Betroffene zu berücksichtigen.

Trotz dieser verschiedenen Zugänge und Perspektiven auf Gestaltungsprozesse stellen beide Ansätze informative Perspektiven für die MZKG respektive HCAI bereit. CTA vertritt mehr eine Außenperspektive zu Gestaltungsfragen, während VSD eine in HCI und somit HCAI deutlich vertrautere Innenperspektive vertritt. Fragestellungen zur MZKG berühren hingegen viele Herausforderungen deren Bewältigung es erfordert, beide Perspektiven zu vereinen, bzw. einen gezielten Perspektivwechsel zwischen ihnen durchzuführen. Die sehr komplexen Fragen zu gesellschaftlichen Auswirkungen von KI-Systemen nehmen in HCAI zunehmend mehr Raum ein, die eine angemessen hohe Sorgfalt und möglichst unvoreingenommene Zuwendung erfordern. Beides sind Charakteristika, die in CTA und VSD aufgrund ihrer Technologieoffenheit durchgehend vertreten sind und die eindeutig zu den Stärken beider Ansätze zählen. Somit besitzen beide Ansätze ein grundsätzliches Potenzial, die in Bezug auf KI-Systeme geäußerten Bedenken zu gesellschaftlichen und ethischen Entwicklungen zu adressieren und somit einen Bedarf in HCAI zu decken.

#### 4.4.2 Menschorientierung

Beide Ansätze werden von dem grundlegenden Bestreben geleitet, menschliche Bedürfnisse und Eigenschaften stärker in den Fokus von Gestaltungsprozessen zu rücken, um auf ihre Weise einen Beitrag zu einer menschorientierten Gestaltung technischer Systeme zu leisten. Wie auch die MZG kommt in beiden Ansätzen verschiedentlich zum Ausdruck, dass die Auseinandersetzung mit menschlichem Leben und menschlichen Werten eine hohe Sorgfalt und Besonnenheit erfordert, besonders um voreiligen, unbedachten Schlüssen und wenig informierten Entscheidungen im Gestaltungsgeschehen vorzubeugen.

Bei VSD wird dies bereits im Namen deutlich, denn das „sensitive“ in VSD adressiert einerseits diesen Umstand und deutet gleichzeitig das Versprechen an, VSD-Anwender mit dem dafür notwendigen Sensibilitätsvermögen auszustatten. In eine ähnliche Richtung weist das „constructive“ in CTA, auch wenn hier die Förderung der Menschorientierung mehr als konstruktiver Beitrag interpretiert wird, um die Gestaltung sozial wünschenswerter technischer Systeme zu fördern.

Auch wird in beiden Ansätze viel Aufwand betrieben, um aktiv in Kontakt mit realen Akteuren und Stakeholdern aus dem Feld zu treten, um deren Ziele, Erfahrungen, Bedenken usw. in Erfahrung zu bringen. Dadurch wird in beiden Ansätzen viel Energie darauf verwendet, ein möglichst ganzheitliches Bild über die Ziele und Eigenschaften der Stakeholder im Zusammenhang mit dem Nutzungskontext zu erhalten, sowie die mit dem Nutzungskontext zusammenhängenden und aus ihm resultierenden Mensch-Technik-Verhältnisse nachzuvollziehen. In CTA und VSD werden Gestaltungsprozesse als sozial eingebettete, wechselwirkende bzw. ko-evolvierende Prozesse betrachtet, in denen es nicht ausreicht, ausschließlich technische Fragen und Aspekte zu analysieren. Beide Ansätze erweitern die in den Technikwissenschaften sicherlich notwendige Vorgehensweise, Gestaltungsprozesse in einer Art künstlichem Vakuum zu betrachten, indem versucht wird, die sozialen Implikationen von Gestaltungsprozessen durch ihre Einordnung in soziale Kontexte zu ergründen. In CTA und VSD wird nicht in einer Gestalterperspektive verblieben, sondern Erkenntnisse werden auch für und mit den involvierten und betroffenen Stakeholdern geteilt, um bei sämtlichen Beteiligten ein hohes Maß an Transparenz zu den gegebenen Zuständen und ablaufenden Vorgängen zu erwirken. Da dies häufig durch aktive Partizipation der Stakeholder selbst erfolgt, kann argumentiert werden, dass CTA und VSD auf ihre Weise den beiden Prinzipien der MZG zur Erfassung des Nutzungskontexts und der Beteiligung von Stakeholdern gerecht werden. Auch werden beide Ansätze dem MZG-Prinzip der Interdisziplinarität in Gestaltungsteams gerecht, da sie zur Formulierung ihrer theoretischen Standpunkte auf ein hohes Maß an interdisziplinärem Wissen zurückgreifen und das Vorhandensein interdisziplinärer Perspektiven in Gestaltungsprozessen begrüßen. Somit verbleibt die Frage, inwieweit die beiden Ansätze dem letzten verbleibende MZG-Prinzip gerecht werden, das die iterative Vorgehensweise in Gestaltungsprozessen vorsieht. Hierzu werden weder in der VSD- noch in der CTA-Literatur sehr viele explizite Aussagen getätigt, ein zyklisches Vorgehen bei der Gestaltung wird nicht konkret gefordert, jedoch auch nicht ausgeschlossen. Eine iterative Anwendung der vorgeschlagenen Methoden bzw. ihre Integration in iterative Frameworks ist jedoch durchaus vorstellbar. Die tripartite methodology von VSD lässt für solche Anpassungen ausreichend Spielraum und auch hinsichtlich CTA ist vorstellbar, dass beispielsweise die Erkenntnisse eines ersten Workshops in die Überarbeitung der zuvor entwickelten soziotechnischen Szenarien einfließt, um zu einem späteren Zeitpunkt erneut Workshops unter Verwendung überarbeiteter Szenarienvarianten durchzuführen. Somit folgen beide Ansätze in hohem Maße den Gestaltungsprinzipien der MZG, wodurch sie ein zentrales Kriterium für genauere Betrachtungen im HCAI-Diskurs erfüllen, auch wenn dort iterative Gestaltungsweisen deutlich stärker im Vordergrund stehen, um Marktveränderungen, Wertverschiebungen oder Prioritätshifts gerecht werden zu können.



Ein weiteres Merkmal der Menschorientierung beider Ansätze ist die Verwendung verschiedener Rollenkonzepte, um unterschiedliche menschliche Eigenschaften und Bedürfnisse in Gestaltungsprozessen zu beschreiben und um unterschiedliche Positionen und Interessensverteilungen zu ermitteln. Dies verdeutlicht eine reflektierte und differenzierte Sichtweise auf menschliches Leben und Handeln im Gestaltungskontext, die für die Zielsetzungen der MZKG notwendig ist.

#### 4.4.3 Methodische Vagheit

Die zuvor erfolgte Darstellung der Methoden beider Ansätze kann den Eindruck erwecken, dass die Literatur zu beiden Ansätzen einen leichten Zugang zu den Methoden gewährt und interessierten Anwendern, ähnlich einer instruktiven Handreichung, eine rasche Einarbeitung und Replikation ermöglicht. Die Literatur beschreibt die Methoden jedoch an vielen entscheidenden Stellen überwiegend unkonkret und vage, sodass auch die zu methodischen Fragestellungen mehrfach referenzierte Literatur nur punktuell leicht nachvollziehbare und belastbare Aussagen zum methodischen Vorgehen beinhalten. Begründet wird diese Vagheit üblicherweise mit dem Hinweis, dass zu konkrete Angaben zum methodischen Vorgehen im Widerspruch stünden zu der Notwendigkeit, die Methodenauswahl an die Besonderheiten eines Gestaltungsprozesses anpassen zu können:

*„While some have critiqued value sensitive design for not being more directive about method [...], we are hesitant to be overly prescriptive. Rather, the challenge is to select a method (or a constellation of integrated methods) that fits a particular project at a particular point in the design process“* (B. Friedman & Hendry, 2019, S. 38)

Diese Argumentation ist schlüssig und nachvollziehbar, besonders angesichts der Tatsache, dass in beiden Ansätzen inhaltlich teilweise hochkomplexe Folge- und Wechselwirkungen sozialer, ethischer und technischer Aspekte im Vordergrund stehen und deshalb Wissen um die Tragweiten und möglichen Konsequenzen von zu rigide formulierten Gestaltungsanweisungen vorhanden ist. Allerdings überrascht diese Argumentation angesichts der Tatsache, dass es sich bei beiden Ansätzen gemäß dem in der Literatur dargelegten Selbstverständnis um Gestaltungsansätze handelt, die die Zielsetzung verfolgen, in möglichst konkreten Gestaltungssituationen Anwendung zu finden. Dies erfordert, dass methodische Vorgehensweisen möglichst transparent, detailliert und nachvollziehbar vermittelt werden, um eine größtmögliche Chance auf ihre Anwendung zu erreichen und somit zu ihrer Robustheit beizutragen. Dem Kriterium eines niederschweligen Zugangs zu Methoden und damit zusammenhängend einer leichten Replizierbarkeit werden beide Ansätze in 2024 jedoch nur punktuell gerecht.

Da eine solche Präzisierung an entscheidenden Stellen überwiegend nicht erfolgt, fällt diese den interessierten Anwendern der CTA- und VSD-Methoden zu. Damit wird erfolgsentscheidende Transferarbeit auf potenzielle Anwender übertragen, womit ein gewisses Risiko verbunden ist. Sollten potenzielle Anwender beispielsweise über wenig Hintergrundwissen zu den Ansätzen verfügen, in der fächerübergreifenden Zusammenar-

beit nur wenig Erfahrung besitzen, oder bislang kaum in Berührung mit dem für die Gestaltung notwendigen, offenen Mindset gekommen sein, besteht die Möglichkeit, bei der Auslegung der Methoden die mit ihnen verfolgten, originären Intentionen der Ansätze fehlzuinterpretieren, ungünstig zu gewichten oder im äußersten Fall ihnen gar entgegenzuwirken. Auch senkt die didaktische Intransparenz die Motivation, tatsächlich mit den Methoden zu arbeiten, da mit der zur Erbringung der Transferleistung notwendigen Einarbeitung ein hoher Ressourcenaufwand verbunden ist, der besonders in Kontexten der Gestaltungspraxis selten aufgebracht werden kann. Insofern ist erkennbar, dass trotz dem Anspruch der Praxistauglichkeit, sich beide Ansätze primär an akademische Zielgruppen richten und primär zum dortigen Wissenserwerb einen Beitrag leisten möchten.

#### 4.4.4 Erfolgskritische Aspekte der CTA-Methodik

In der Methodik beider Ansätze wird auf bewährte Praktiken sozialwissenschaftlicher Forschung zurückgegriffen. In VSD werden hier vergleichsweise mehr Möglichkeiten ausgeschöpft und auch die potenziellen Vorteile der Kombination qualitativer und quantitativer Methoden hervorgehoben, während in CTA neben Feldstudien die Kombination aus Szenarien und Workshops die augenscheinlich präferierten Werkzeuge darstellt. Durch diese stärkere methodische Fokussierung haben die Methoden in CTA einen vergleichsweise hohen Anteil daran, den CTA-Interventionen Erfolg zu gewährleisten. Die sozio-technischen Szenarien fallen unter diesem Aspekt besonders ins Gewicht, da sie den durchgehenden, inhaltlichen Bezugspunkt in den Workshops darstellen. Sie erfüllen in CTA eine maßgebliche Transferfunktion, da sie (1) den erfassten Status quo des Technologiefelds in für die Akteure ausreichend wiedererkennbarem Maße wiedergeben müssen und (2) mögliche Entwicklungen skizzieren müssen, die für die Workshopteilnehmer ein Mindestmaß an Bedeutung besitzen. Sollten die Szenarien beispielsweise zu vage formuliert sein oder an Dynamiken des Technologiefelds ansetzen, deren Bedeutung für die Akteure nicht ausreichend überzeugend dargestellt werden kann, sinken die Chancen erheblich, im Rahmen der Workshops zielführende Diskussionen zu ermöglichen und somit den CTA-Interventionen einen Erfolg zu bescheiden. Die konkrete Form der Szenarien, die verwendete Sprache und die Vorgehensweise zur Auswahl der Inhalte ist somit maßgeblich entscheidend, um die Erfolgchancen von CTA zu steigern und rückblickend beurteilen zu können. Angesichts dieser kritischen Bedeutung ist es verwunderlich, dass die CTA-Literatur, mit nur wenigen Ausnahmen (Parandian, 2012; Parandian & Rip, 2013), kaum belastbare Angaben zur Entwicklung der Szenarien macht und nur wenig konkrete Beispiele für in CTA-Projekten entwickelte Szenarien bereitstellt. Auch finden sich kaum Schilderungen zu Erfahrungswerten oder bewährten Praktiken, um den Bedeutungsgehalt der Szenarien abzusichern, wodurch an dieser Stelle ein hohes Potenzial weitgehend ungenutzt bleibt, interessierten CTA-Anwendern wertvolle Erkenntnisse zu vermitteln. Dies überrascht angesichts der langjährigen Projekterfahrung, die in der CTA-Literatur konzentriert ist und in der das Erfahrungswissen zur zentralen Funktion der Szenarien verankert ist, Akteure zur Teilnahme an den Workshops zu motivieren und zu überzeugen. Diesem ungenutzten Potenzial verleiht auch der Aspekt zusätzliches Gewicht,

dass die Szenariotechnik an sich eine nur geringfügig standardisierte Methode darstellt, sodass es fraglich ist, inwieweit interessierte CTA-Anwender bei der Suche nach zusätzlichem Wissen zur Szenarioentwicklung außerhalb der CTA-Literatur eine bedeutsame Wissenslücke schließen können.

#### 4.4.5 Operationalisierbarkeit für die KI-Gestaltung

Beide Ansätze sind hinsichtlich ihrer Anwendbarkeit mit einer bewusst gewählten Technologie- und Kontextoffenheit konzipiert worden. In CTA steht keine bestimmte Technologie im Vordergrund, die Auswahl bzw. Relevanz einer Technologie die in CTA betrachtet wird, ergibt sich aus ihrer Relevanz für gesellschaftliche Dynamiken auf der Makroebene. VSD wurde mit dem Ansinnen entwickelt, in Gestaltungsprozessen von IT-Systemen Anwendung zu finden, findet jedoch etwa seit den 2010er-Jahren auch Gebrauch in der Gestaltung nicht-informatischer Systeme. Beide Ansätze sind somit nicht dezidiert für eine Anwendung in KI-Gestaltungsprozessen entwickelt worden, besitzen jedoch aufgrund der in ihnen angelegten Offenheit und Flexibilität die Möglichkeit, auf KI als ein Anwendungsfeld angepasst und angewendet zu werden. Diese Technologieoffenheit beider Ansätze lässt viel Spielraum zur Konkretisierung und Anpassung auf die Besonderheiten von KIGP, wobei die traditionelle Nähe zur Informatik VSD im Vergleich mit CTA für die Debattenkultur in HCAI tendenziell leichter zugänglich und leichter adaptierbar macht. Die grundsätzliche methodische Flexibilität beider Ansätze ist den Zielen der MZKG insofern zuträglich, dass sie gewährleistet, den individuellen Gestaltungskontexten und den im Fokus stehenden konkreten KI-Anwendungen begegnen zu können. Die unpräzise Beschreibung der Methodenanwendung auf Detailebene hemmt jedoch einen direkten Transfer in konkrete Gestaltungsprozesse erheblich, was die Anschlussfähigkeit beider Ansätze für die Gestaltungspraxis absenkt.

## 4.5 Zusammenfassung

Die Inhalte der beiden Ansätze CTA und VSD eint die Bestrebung, menschliche Anforderungen und Bedürfnisse in den Mittelpunkt von Gestaltungsprozessen zu rücken. Ihre Motivationen für die Realisierung dieser Bestrebung stammen aus verschiedenen Denkschulen, weshalb in beiden Ansätzen leicht verschiedene Konzeptionen menschlichen Lebens und Handelns in Gestaltungsprozessen zum Tragen kommen und auch die vorgeschlagenen Methoden zur Umsetzung der Menschorientierung verschiedene Strategien aufgreifen und einsetzen. In CTA liegt der Fokus darauf, Gestaltungsprozesse in Form von „bridging events“ (Rip, 2018, S. 102) auf gesellschaftliche Anforderungen auszurichten, während in VSD die Auseinandersetzung mit Werten und Wertekonflikten im Vordergrund steht, die mithilfe der tripartite methodology aus konzeptionellen, empirischen und technischen Perspektiven untersucht werden. Beide Ansätze können zur Gestaltung von KI-Systemen eingesetzt werden und sind darüber hinaus kompatibel zur Gestaltungsphilosophie der MZG, weshalb sie sich grundsätzlich für eine Erwägung zu Fragestellungen der MZKG in HCAI eignen. Hierbei gilt es jedoch zu berücksichtigen, dass CTA und VSD zu Zeiten entwickelt wurden, in welchen der Stand der Technik von KI nicht dem der Gegenwart entsprach und auch, dass beide Ansätze nicht mit dem Ansinnen entwickelt wurden, den spezifischen Charakteristika von KIGP begegnen zu können. Um genauere Aussagen zur Anschlussfähigkeit beider Ansätze zum MZKG-Diskurs treffen zu können ist deshalb eine Auseinandersetzung mit CTA und VSD auf Detailebene erforderlich. Da für den MZKG-Diskurs die in Gestaltungsprozessen verwendeten Konzeptionen zu menschlichem Leben und Handeln einen zentralen Aspekt darstellen, widme ich mich anschließend den Rollenkonzepten in CTA und VSD.

## 5 Rollenkonzepte in CTA und VSD

Wie in Gestaltungsprozessen über Mensch-Technik-Verhältnisse nachgedacht und gesprochen wird, hat entscheidende Auswirkungen auf die Kooperationsanstrengungen aller am Prozess Beteiligten und auch auf die erwartbaren Resultate des Prozesses. Der Grundstein für die Umgangsformen in und Ergebnisse von Gestaltungsprozessen werden häufig in den in ihnen verwendeten Gestaltungsphilosophien und den daraus hervorgegangenen Methoden gelegt. Dieser Umstand und die damit verbundenen Herausforderungen und Implikationen stehen gegenwärtig auch im Mittelpunkt der Debatten zur MZKG, da dort häufig unklar ist, auf welche Vorstellungen des Menschen sich unter dem Label der Menschzentriertheit berufen wird. Dies macht es für Communities wie der HCAI erforderlich, sich mit den in KIGP verwendeten Konzeptionen menschlichen Lebens und Handelns auseinanderzusetzen, insbesondere vor dem Hintergrund der mit KI-Systemen verbundenen weitreichenden sozialen und ethischen Implikationen. Um die konzeptionelle Anschlussfähigkeit von CTA und VSD zum MZKG-Diskurs einschätzen und beurteilen zu können, ist daher ein umfassendes Verständnis über die in den beiden Ansätzen verwendeten Konzeptionen menschlichen Lebens und Handelns notwendig. Der Untersuchung solcher Verständnisse zu Rollen von Menschen in Gestaltungsprozessen ist dieses Kapitel gewidmet. Zu Beginn werde ich definieren, welches Verständnis von Rollenkonzept der Untersuchung zugrunde liegt (siehe Kap. 5.2), um nach der Beschreibung meiner Vorgehensweise (siehe Kap. 5.3) die in CTA und VSD verwendeten Rollenkonzepte vorzustellen (Forschungsfrage 1; siehe Kap. 5.4). Daran schließt die Untersuchung der Ursprünge der Rollenkonzepte an (Forschungsfrage 2; siehe Kap. 5.5), auf die die Darstellung der blinden Flecke der Rollenkonzepte folgt (Forschungsfrage 3; siehe Kap. 5.6). Die daraus resultierenden Erkenntnisse diskutiere ich anschließend vor dem Hintergrund der übergeordneten Forschungsfrage der Dissertation (siehe Kap. 5.7).

### 5.1 Einführung

Die ihm zugrunde gelegten Annahmen zu menschlichem Leben und Handeln haben entscheidende Auswirkungen auf den Verlauf und die Resultate eines Gestaltungsprozesses. Sie dienen der Definition von Anforderungen an das zu gestaltende System, sie haben maßgeblichen Einfluss auf die im Gestaltungsprozess verwendete Sprache (Bendoukha, 2009) und sie prägen in hohem Maße die Einschätzungen zu den übergeordneten Mensch-Technik-Verhältnissen, die im Rahmen des Gestaltungsprozesses als relevant und gültig erachtet werden. So unterscheidet sich beispielsweise ein Gestaltungsprozess, in dem Menschen als überwiegend kompetent im Umgang mit Technik betrachtet werden, deutlich von einem Gestaltungsprozess, dem die Vorstellung eines Dumbest Assumable Users (Spiekermann, 2010) oder die eines notorischen Technikvermeiders als eine zentrale Prämisse dient. Derartige Rollenkonzepte stellen Modelle dar, die bestimmte Annahmen über menschliche Eigenschaften und Fähigkeiten bündeln und sie Menschengruppen im Rahmen eines Gestaltungsprozesses zuschreiben. Sie sind ferner das Ergebnis von Abstrak-



tionsvorgängen, deren Systematik und Reflexionsgrad stark variieren kann, deren Durchführung jedoch in sämtlichen Gestaltungskontexten als notwendig erachtet wird und daher eine gängige Praxis darstellt. Die Notwendigkeit der Durchführung solcher Abstraktionsvorgänge ist bedingt durch die verschiedenen Formen und hohen Grade von Komplexität, die in Gestaltungskontexten vorherrschen bzw. aufeinandertreffen, und die gestaltende Disziplinen entsprechend zu balancieren und zu reduzieren suchen. Auf technischer Ebene stehen in Gestaltungsprozessen inhaltlich komplexe Sachverhalte und Verfahren im Vordergrund, wie etwa die mathematischen und statistischen Verfahren in der KI-Gestaltung. Die solchen Verfahren zugrundeliegenden Gesetzmäßigkeiten sollen technisch operationalisiert werden, um sie für den Menschen nutzbar zu machen und um seine Handlungsspielräume zu erweitern. Auf organisatorischer Ebene handelt es sich bei Gestaltungsprozessen um hochdynamische Unternehmungen, da sie angesichts begrenzter zeitlicher Ressourcen eine durchdachte Koordination und ergebnisorientierte Zuteilung finanzieller und personeller Ressourcen erfordert. Und auch auf sozialer Ebene begegnen Gestaltungsprozesse einer stark ausgeprägten Form von Komplexität, denn die gestalteten Systeme sollen sich möglichst konfliktfrei in verschiedene soziale Settings integrieren lassen und dort eine für Menschen einfache und unbedenkliche Nutzung ermöglichen.

Um eine Aussicht auf gestalterischen Erfolg zu wahren, müssen in Gestaltungsprozessen auf diesen verschiedenen Ebenen Entscheidungen herbeigeführt und Maßnahmen getroffen werden, die eine Kompromissorientierung und Komplexitätsreduktion zum Ziel und Inhalt haben. Zur Herbeiführung dieser werden Gestaltungsansätze, Managementframeworks, Leitlinien und weitere Entscheidungshilfen angewandt, die bestimmte Aspekte dieser Komplexitätsebenen in den Blick nehmen und die ihrerseits verschiedene Mittel und Strategien kombinieren. Die Abstraktion erfährt unter diesen Strategien eine besondere Bedeutung, da mit ihrer Hilfe die Komplexität aller drei Ebenen an entscheidenden Stellen reduziert werden kann. Rollenkonzepte stellen eine Ergebnisform von Abstraktionsarbeit dar, die als Modelle menschlicher Lebens- und Handelsweisen besonders die soziale und organisatorische Komplexität eines Gestaltungsprozesses reduzieren helfen. Die komplexitätsreduzierende Funktion resultiert dabei aus der Summe verschiedener Teilfunktionen: (1) Rollenkonzepte ermöglichen die Identifikation und Beschreibung der an der Gestaltung beteiligten und von ihren Resultaten betroffenen Stakeholdergruppen und (2) sie erlauben, die tieferliegenden Motivationen und Zielsetzungen der Stakeholder zu ergründen, nachzuvollziehen und zu antizipieren. Durch die in (1) und (2) bereitgestellten Erkenntnisse (3) unterstützen sie zudem eine durch relevante Informationen gestützte Herbeiführung von Entscheidungen zur konkreten Ausgestaltung eines technischen Systems und der dafür erforderlichen Ressourcenverteilung. Rollenkonzepte stellen durchgängige Orientierungsgrößen für den Gestaltungsprozess dar, die die relevanten Ausschnitte sozialer Realität einzugrenzen helfen und dabei unterstützen, diesen Fokus beizubehalten, zu erweitern oder zu präzisieren. Durch ihre Entscheidungsunterstützungsfunktion erfahren sie zudem eine Bedeutung für die Evaluation von Gestaltungsaktivitäten. In Summe vereinen Rollenkonzepte die drei für abstrakte Modelle typischen Funktionen der Beschreibung und Erklärung von Sachverhalten sowie der Herbeiführung von Entscheidungen (Strahnger, 2016).

Eine weitere bedeutsame Funktion von Rollenkonzepten besteht darin, die Menschorientierung in Gestaltungsprozessen zu gewährleisten und zu steuern. Sie sind ein entscheidender Bestandteil der Art und Weise, wie menschliche Lebens- und Handlungsweisen in Gestaltungsprozessen repräsentiert und adressiert werden, da sie neben rein deskriptiven häufig auch wertende Positionen und Einschätzungen beinhalten. Dies ist einer der Gründe, weshalb ihrer Ausarbeitung und ihrem Gebrauch in menschorientierten Gestaltungsansätzen wie der MZG, und den mit solchen Ansätzen arbeitenden Fachbereichen wie der HCAI, eine hohe Aufmerksamkeit gewidmet wird. In Verbindung mit zyklisch ausgerichteten Gestaltungsprozessen kann zudem eine kontinuierliche Überprüfung ihrer Angemessenheit gewährleistet werden. Mit Blick auf die Fragestellung der Dissertation stellt das Element der Rollenkonzepte somit ein geeignetes Merkmal dar, um die Anschlussfähigkeit von CTA und VSD zu HCAI zu untersuchen.

## 5.2 Rollenverständnisse im Gestaltungskontext

In der Auseinandersetzung mit Rollenkonzepten ist der Rollenbegriff von zentraler Bedeutung, der jedoch im allgemeinen wie auch akademischen Sprachgebrauch vielfältig besetzt ist. Dies zeigt bereits die etymologische Annäherung an den Rollenbegriff, die zunächst zum lateinischen Wort „rota“ führt. Rota hat die Bedeutung eines Rades, einer Walze oder auch einer rollenden Scheibe, die im deutschen und französischen Sprachgebrauch des 15. Jahrhunderts auf die Pergament- und Schriftrolle übertragen wird. Im 16. Jahrhundert überträgt sie sich weiter auf den Lebensbereich der Schauspielkunst, wo zur damaligen Zeit Papierstreifen auf Rollen dazu benutzt werden, die Texte und Handlungen von Schauspielern eines Theaterstücks festzuhalten. Ausgehend von diesem Kontext wandelt sich ab Mitte des 17. Jahrhunderts die Bedeutung hin zur Bedeutung einer (sozialen) Position, Funktion oder Aufgabe, auf die die bis heute erhaltene und oftmals mehrdeutige Redewendung „eine Rolle spielen“ zurückgeführt wird. Von Interesse für die nachfolgenden Ausführungen ist die metaphorische Funktion dieser Redewendung, die auch im wissenschaftlichen Diskurs unter der Referenz der Theater- oder Bühnenmetapher (König, 2019) zur Analyse und Beschreibung sozialer Interaktionen verwendet wird. Diese Verwendung beginnt sich im 20. Jahrhundert im Diskurs um die „Denkfigur“ Rolle (Schwinger, 2007) zu entspinnen. Verschiedene Untersuchungen der Kulturanthropologie, Soziologie und Sozialpsychologie beginnen über den Rollenbegriff zu reflektieren, den sie bislang dazu verwenden, Wechselwirkungen zwischen Individuum und Gesellschaft zu beschreiben, sowie um Erklärungen für die Zusammenhänge und Wechselwirkungen sozialen Handelns zu formulieren. Die über die Bedeutung und Verwendung des Rollenbegriffs reflektierenden Arbeiten bilden das Fundament des heute reichhaltigen Wissenskorpus zur Rollentheorie, die sich systematisch mit dem Rollenbegriff auseinandersetzt. In rollentheoretischen Untersuchungen wird eine soziale Rolle häufig definiert als

*„Bündel normativer Verhaltenserwartungen, die von einer Bezugsgruppe oder mehreren Bezugsgruppen an Inhaber bestimmter sozialer Positionen herangebracht werden“ (Peuckert, 1992, S. 252).*

Eine soziale Rolle sorgt dadurch

*„für regelmäßiges, vorhersagbares Verhalten als Voraussetzung für kontinuierlich planbare Interaktionen und erfüll[t] somit eine allgemeine soziale Orientierungsfunktion“ (Peuckert, 1992, S. 252).*

Soziale Rollen können demnach Menschengruppen von anderen Menschengruppen zugewiesen oder zugeschrieben werden, um basierend auf diesen Zuschreibungen Handlungen zu planen oder durchzuführen. Diese Interpretation der sozialen Rolle motiviert mich für die Orientierung an der Rollentheorie zur Entwicklung einer Arbeitsdefinition und eines Metamodells zum Konstrukt Rollenkonzept für die Dissertation, da sie Aspekte beinhaltet, die auch für die Verwendung in Gestaltungskontexten bedeutsam sind. Zusätzlich verspricht die Verortung der Rollentheorie in geistes- und sozialwissenschaftlichen Diskursen eine angemessen reflektierte Auseinandersetzung mit dem Untersuchungsgegenstand menschlichen Lebens und Handelns, die auch für menschenzentrierte Gestaltungsprozesse von zentraler Bedeutung sind.

Von den rollentheoretischen Betrachtungen sind besonders jene von Interesse, die sich den Beziehungen zwischen Individuum und Gesellschaft auf einer im sozialwissenschaftlichen Duktus Mikro- oder Mesoebene annähern. Sie erlauben die Betrachtung menschlichen Lebens und Handelns auf der Ebene von Gruppen sowie auf der Ebene von Individuen, und etablieren dadurch einen Spielraum für die Durchführung von Perspektivwechseln und von Abstandsverringerungen in der Auseinandersetzung mit menschlichen Ansichten und Handlungsweisen. Unter diesen Gesichtspunkten tritt besonders das von dem Kulturanthropologen Ralph Linton formulierte Rollenkonzept hervor, mit dessen Hilfe einige zentrale Begriffe und Begriffszusammenhänge erarbeitet werden können. Ein anschließender Blick in die Modellierung von Softwareprozessen ermöglicht es mir nachzuzeichnen, wie Rollenkonzepte in Gestaltungskontexte eingeordnet werden können, und welche Unterschiede zwischen einer mehr sozialorientierten und mehr prozessorientierten Interpretation und Verwendung von Rollenkonzepten bestehen. Die Zusammenführung beider Perspektiven mündet schließlich in ein integratives Modell von Rollenkonzept, das meine nachfolgende Analyse der Rollenkonzepte in CTA und VSD anleitet. Zur visuellen Darstellung der verschiedenen Rollenkonzepte greife ich auf die Darstellungsform von Entity-Relationship-Modellen in der Chen-Notation (P. Chen, 1976) zurück, die sich in vergleichbaren Zusammenhängen zur Visualisierung von Rollenkonzepten als geeignet erwiesen hat (Opitz, 2009).



### 5.2.1 Rollenverständnis von Ralph Linton

In der Rollentheorie gilt der Kulturanthropologe Ralph Linton als der erste Forscher, der sich mit dem Rollenbegriff auseinandersetzt und erste Vorschläge zu einer Taxonomie ausarbeitet. Maßgeblich dafür ist seine einschlägige Arbeit „A study of man“ (Linton, 1936), in der er sich mit dem Beziehungsgeflecht von Individuum und Gesellschaft auseinandersetzt und die dieses Beziehungsgeflecht auszeichnenden Wechselwirkungen zu ordnen versucht. Gesellschaften bestehen für Linton aus Gruppen von Individuen, die zusammen leben und arbeiten und die ihre Ansichten und Verhaltensweisen konstant adaptieren (Linton, 1936, S. 253). Dies stellt die grundlegende Struktur von Gesellschaften dar, die durch die Handlungen und Lebensweisen von Individuen permanent aktiv ist. Parallel zu dieser Struktur bestehen im Lintonschen Denken zusätzlich die abstrakten Gebilde der Sozialen Systeme, die von Individuen nur zeitweise aktiviert und gelebt werden, und die von Individuen losgelöste und unabhängige Anordnungen verschiedener Stati und Rollen darstellen (Linton, 1936, S. 253). Jedes Individuum kann zeitgleich mehrere Stati besitzen, da es üblicherweise an einer Reihe verschiedener sozialer Prozesse (social patterns) beteiligt ist. Linton unterscheidet darüber hinaus verschiedene Arten von Stati. *Ein* Status steht für ein unabhängig von Individuen bestehendes gedankliches Konstrukt, „a position in a particular pattern“ (Linton, 1936, S. 113). *Der* Status verkörpert hingegen die Gesamtheit sämtlicher Stati, die ein Individuum zu einem bestimmten Zeitpunkt besitzt. Linton illustriert den Unterschied an dem Beispiel eines fiktiven Mr. Jones, dessen (Gesamt-)Status sich durch verschiedene (einzelne) Stati wie Bürger, Anwalt oder Ehemann von Mrs. Jones zusammensetzt. Ferner unterscheidet Linton zwischen zugeschriebenen (ascribed) und erworbenen (achieved) Stati, wobei zugeschriebene Stati die von Individuen bei der Geburt erhaltenen Merkmale wie beispielsweise Geschlecht, Rasse oder Familienstand bezeichnen, während erworbene Stati Individuen sich durch das Aufbringen diverser Anstrengungen selbst erwerben. Ein jeder Status umfasst zudem bestimmte Rechte und Pflichten, die jedoch erst dann ersichtlich werden, wenn ein Individuum einen bestimmten Status einnimmt. Aus diesem Grund stellt es nach Ansicht von Linton eine Schwierigkeit dar, bei der abstrakten Untersuchung menschlichen Handelns und sozialer Strukturen eine Unterscheidung zwischen den Stati und den diese Stati auszeichnenden Rechte und Pflichten durchzuführen.

Eine Rolle stellt für Linton nun den dynamischen Aspekt eines Status dar, die „sum total of the culture patterns associated with a particular status“ (Linton, 1952, S. 50). Setzt ein Individuum die mit einem Status zusammenhängenden Rechte und Pflichten in die Tat um, übt es eine Rolle aus. Rolle und Status sind deshalb für Linton untrennbar miteinander verbunden: „There are no roles without statuses or statuses without roles“ (Linton, 1936, S. 114). Ähnlich den Stati kann ein Individuum mehrere Rollen besitzen bzw. einnehmen, und wie bei den Stati unterscheidet Linton auch hier zwischen *einer* Rolle und *der* Rolle. *Die* Rolle beschreibt dabei sämtliche Aktivitäten und Funktionen, in denen ein Individuum zum gesellschaftlichen Leben beiträgt und was es von seiner Gesellschaft als Gegenleistung für das Ausüben seiner Rollen erwarten kann. Eine Rolle umfasst zudem Einstellungen, Wertvorstellungen und Verhaltensweisen eines Individuums, sowie die Erwartungen von Individuen weiterer Stati im selben sozialen System (Linton,

1952, S. 50). Äußerlich betrachtet ist jede Rolle mit einem bestimmten Status verknüpft, für ein Individuum selbst besteht zwischen den beiden Konstrukten Rolle und Status jedoch ein bedeutsamer Unterschied. Dies führt Linton darauf zurück, dass Stati einem Individuum üblicherweise von anderen Mitgliedern einer Gesellschaft zugewiesen oder zugeschrieben werden, während ein Individuum sich überwiegend selbst Rollen aneignet, um seinen gegenwärtigen oder für die Zukunft antizipierten Stati gerecht zu werden.

In Summe stellen im Lintonschen Rollenverständnis die Kombinationen sämtlicher Stati und Rollen das Minimum an Einstellungen und Verhaltensweisen dar, die Individuen annehmen müssen, um am Sozialgeschehen teilhaben zu können und an dessen Weiterentwicklung mitwirken zu können. Status und Rolle dienen dazu, die für Individuen idealen Muster (ideal patterns) ihres sozialen Lebens zu identifizieren und auf die für ihre Lebenssituation wesentlichen bzw. hilfreichen Muster zu reduzieren. Sie unterstützen Individuen bei der Auswahl von Einstellungen und der Organisation von Verhaltensweisen im Sozialgeschehen, sodass sich die von ihnen ausgewählten sozialen Muster mit den Mustern anderer Individuen vertragen, die am Sozialgeschehen teilhaben und mitwirken. Eine Skizze des Rollenkonzeptverständnisses von Ralph Linton habe ich in einem Entity-Relationship-Modell zusammengefasst (siehe Abb. 38)

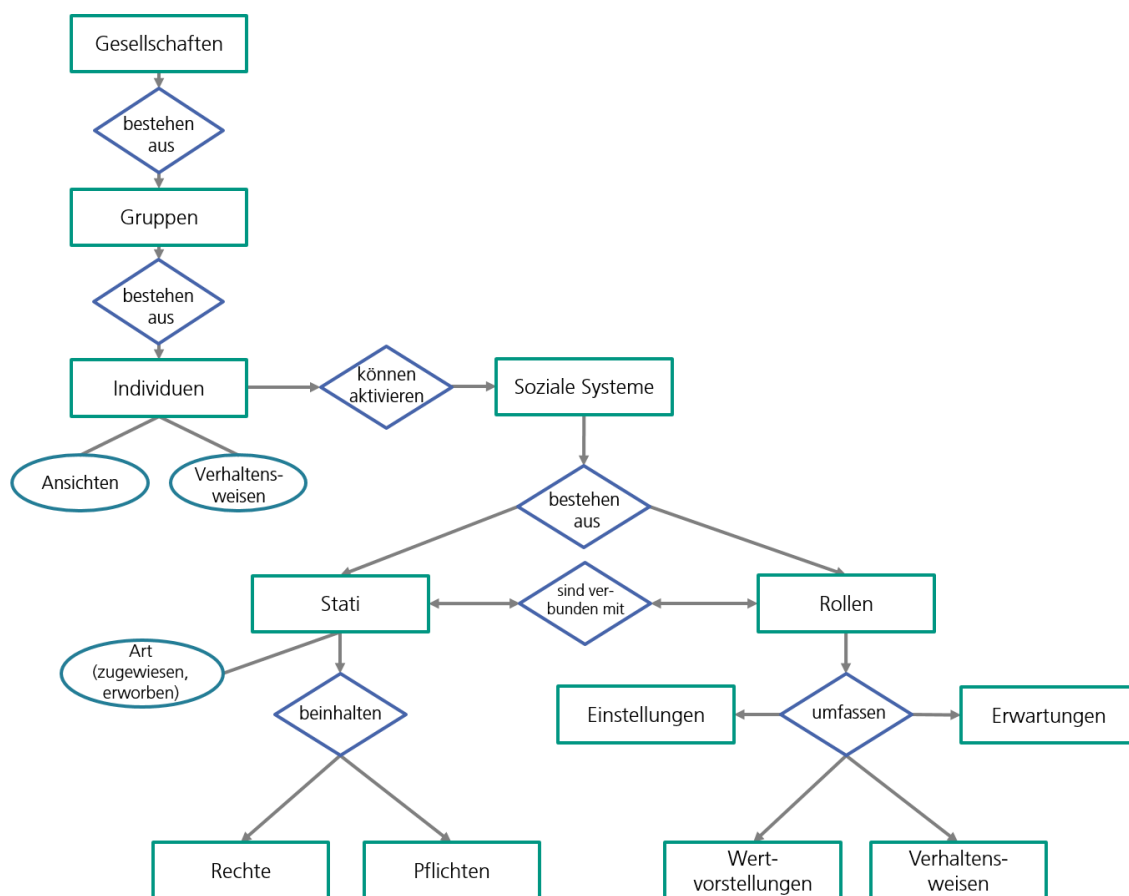


Abbildung 38: Entity-Relationship-Modell des Rollenverständnisses von Ralph Linton

## 5.2.2 Rollenverständnis in der Modellierung von Softwareentwicklungsprozessen

In der Modellierung von Softwareentwicklungsprozessen (MSEP) werden Rollenkonzepte verwendet, die auf inhaltlicher Ebene Parallelen zum Rollenverständnis von Ralph Linton aufweisen, deren Verwendungszweck sich jedoch deutlich von der Untersuchung sozialer Interaktionen und Wechselwirkungen unterscheidet. Rollenkonzepte werden in der MSEP ebenfalls dazu verwendet, menschliche Handlungen zu beschreiben, dies geschieht hier jedoch primär vor dem Hintergrund, die Qualität der zu entwickelnden Software zu gewährleisten und die Zusammenarbeit der an der Entwicklung beteiligten Personen zu unterstützen. Mit der Betrachtung von Softwareentwicklungsprozessen als einem Ausschnitt sozialen Geschehens ist der Fokus hier ein deutlich engerer und hat entsprechend einen deutlicher abgegrenzten Gültigkeitsbereich wie rollentheoretische Betrachtungen und Überlegungen.

Die MSEP kennt verschiedene Ansätze und Verfahren, in denen mit der Denkfigur Rolle gearbeitet wird, um die verschiedenen Aufgaben und Zuständigkeiten beteiligter Expertisen zu unterscheiden und mit Blick auf ein zu gestaltendes Softwareprodukt zu koordinieren. Anders als in der Rollentheorie wird in Arbeiten der MSEP jedoch weniger über die Verwendung des Rollenbegriffs und den damit verbundenen Implikationen reflektiert. Die Beschreibung menschlichen Verhaltens bzw. die Zuweisung von Aufgaben und Tätigkeiten mithilfe von Rollenkonzepten hat sich dort als gängige Praxis etabliert. Entsprechend gehen Arbeiten der MSEP deutlich weniger ins Detail, wenn sie sich mit den ihren Modellierungsverfahren und -ergebnissen zugrundeliegenden Rollenverständnissen auseinandersetzen. Eine Arbeit, die eine ausreichende Detailebene erreicht und in der auch Bezüge zum Rollenkonzept von Ralph Linton ersichtlich werden, ist das bei Acuña et al. (Acuña et al., 2001) dargelegte Prozessmodell (siehe Abb. 39), das mit den Elementen Actor, Role, Activity und Product die zentralen Elemente eines Softwareentwicklungsprozessmodells darstellt und ihre Zusammenhänge erläutert.

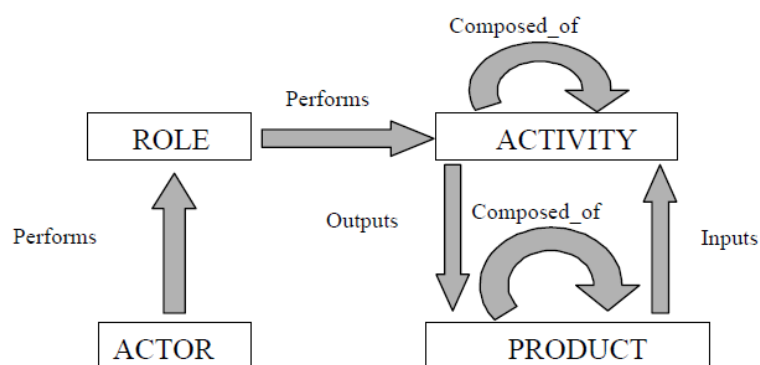


Abbildung 39: Elemente eines Prozessmodells in der MSEP nach Acuña et al. (2001, S. 199)

Unter einem Actor wird hier eine „entity that executes a process“ (Acuña et al., 2001, S. 198) verstanden. Eine solche Entität kann eine von zwei Ausprägungen besitzen, sie kann einen menschlichen Akteur repräsentieren oder eine technische Komponente eines Softwaresystems. Charakterisiert wird ein Actor durch die Eigenschaften der Rollen, die er im Rahmen des Softwareentwicklungsprozesses ausübt. Diese stellen eine definierte Menge von Verantwortlichkeiten, Rechten und Pflichten dar, die ein Actor benötigt, um Aktivitäten im Softwareentwicklungsprozess durchzuführen. Analog zum Rollenverständnis von Linton kann auch ein Actor in einem Softwareentwicklungsprozess mehrere Rollen einnehmen bzw. besitzen. Eine Rolle umfasst zudem eine bestimmte Anzahl von Aktivitäten, die definiert sind als Phasen des Entwicklungsprozesses, die von außen erkennbare Zustandsänderungen in einem Softwareprodukt bewirken. Eine Aktivität umfasst Verfahren, Regeln, Richtlinien und Ziele, die der Erzeugung oder Veränderung vorgegebener Eigenschaften eines Softwareprodukts dienen sollen. Eine solche Veränderung kann die konkrete Form eines (technischen) Inputs oder Outputs annehmen, oder auch eine andere Form wie beispielsweise ein Datenbankschema darstellen. Aktivitäten können ferner in Sub-Aktivitäten (elementary activities) unterteilt werden, ähnlich der Unterscheidung *eines* Status und *dem* Status bei Linton. Ein Artefakt oder Produkt repräsentiert das (Zwischen-) Ergebnis des Softwareentwicklungsprozesses. Mit seiner Hilfe können weitere Produkte erschaffen werden und es kann im Laufe weiterer Entwicklungsprozesse zusätzliche Änderungen erfahren. Ein Softwareprodukt, dass an Benutzer übergeben werden soll, kann aus verschiedenen (Teil-)Produkten bestehen (im Modell: „Composed\_of“). Dieses Modell zeigt exemplarisch, dass in der MSEP im Vergleich zur Rollentheorie deutlicher instrumentelle, und deutlicher geschlossene Vorstellungen von Rollenkonzepten zum Tragen kommen. Der Rollenbegriff dient der Unterscheidung von Aufgaben und Zuständigkeiten sowie der Identifikation von benötigten Expertisen und angestrebten Features zur Realisierung eines Softwareprodukts. Auch dieses Rollenkonzeptverständnis habe ich in einem Entity-Relationship-Modell skizzenhaft zusammengefasst (siehe Abb. 40).

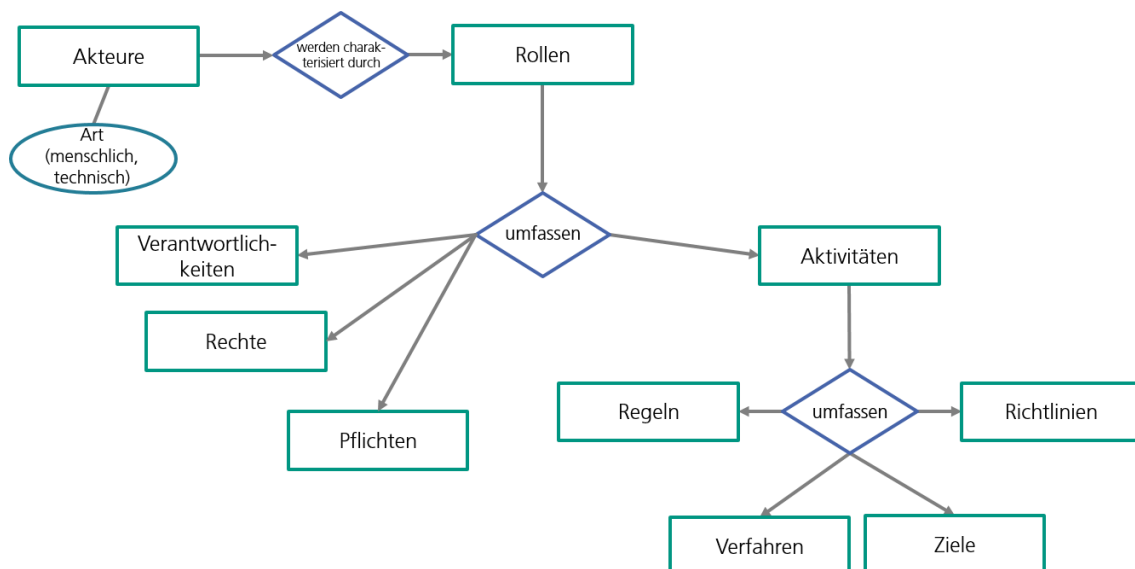


Abbildung 40: Ein in der Softwareprozessmodellierung verwendetes Rollenverständnis

### 5.2.3 Integratives Verständnis von Rollenkonzept

Das von mir angewendete Verständnis von Rollenkonzepten ist bezüglich seines Abstraktionsgrades zwischen den beiden vorgestellten Rollenverständnissen angesiedelt und versucht, Ausschnitte aus beiden miteinander zu verknüpfen. Im Sinne eines Erklärmodells fokussiert es Prozesse zur Gestaltung technischer Systeme als Ausschnitte sozialen Handelns, innerhalb derer die Betrachtung menschlichen Verhaltens auf zwei Beziehungsebenen von Bedeutung ist. Zum einen handelt es sich dabei um die zwischenmenschlichen Beziehungen und Interaktionen, da diese in Form von Kooperations- und Kommunikationsvorgängen Gestaltungsprozesse darstellen. Zudem sind auch die Beziehungen von Menschen zu technischen Systemen von hohem Interesse, da sie den Verlauf von Gestaltungsprozessen maßgeblich prägen und sich in deren Verlauf verändern können. In Gestaltungsprozessen sind diese beiden Beziehungsebenen der zwischenmenschlichen und der Mensch-Technik-Verhältnisse häufig miteinander verwoben, zwischen ihnen bestehen häufig Zusammenhänge und Abhängigkeiten, die durch einen Gestaltungsprozess geschaffen oder entscheidend verändert werden. Eine Unterscheidung zwischen diesen Beziehungsebenen fällt daher nicht immer leicht, kann aber für die Ausrichtung eines Gestaltungsprozesses oder die Evaluation einzelner Gestaltungsaktivitäten wertvoll und notwendig sein – besonders, wenn diese mit der Bezeichnung „mensch-zentriert“ versehen werden sollen.

In dem von mir entwickelten Erklärmodell zur Untersuchung der Rollenkonzepte in CTA und VSD kann die Betrachtung von Menschen grundsätzlich auf zwei Ebenen erfolgen: auf der Ebene von Gruppen, sowie auf der Ebene von Individuen. Die Betrachtung von Gruppen ist dabei die häufiger verwendete, jedoch kann es bei bestimmten Fragestellungen auch zielführend sein, sich punktuell mit den Zielsetzungen oder Verhaltensweisen von Individuen auseinanderzusetzen. Innerhalb von Gestaltungsprozessen üben Gruppen und Individuen verschiedene Rollen aus, von denen sie zur gleichen Zeit mehrere besitzen können. Rollen werden charakterisiert durch verschiedene beschreibende und antizipierende Eigenschaften, die für einen Gestaltungsprozess relevant sind. Bei den mehr beschreibenden Eigenschaften handelt es sich um Aktivitäten, Verantwortlichkeiten, Rechte und Pflichten, die zum zielführenden Fortschritt eines Gestaltungsprozesses benötigt werden. Sie beziehen sich primär auf die technischen und organisatorischen Aspekte von Gestaltungsprozessen, die aufgrund ihrer Notwendigkeit zur Realisierung eines Gestaltungsvorhabens leicht zu identifizieren und zu definieren sind. Sie beschreiben vornehmlich die fachlichen Expertisen sowie die Kooperationszusammenhänge der im Gestaltungsprozess aktiven und involvierten Gruppen und Individuen. Zusätzlich können Rollen Eigenschaften beschreiben, die sich auf die variablen sozialen Aspekte beziehen, die mit der Gestaltung eines technischen Systems adressiert werden sollen. Typische Ausprägungen solcher situativ bedeutsamen Eigenschaften sind Erwartungen, Einstellungen, Wertvorstellungen, Verhaltensweisen, Fähigkeiten und Ziele von Menschen. Diese Eigenschaften beschreiben und beleuchten die mehr sozialen und kulturellen Dynamiken, die in einem Gestaltungsprozess zum Tragen kommen und die dieser berührt. Über diesen Typus von Eigenschaften können im Vorfeld häufig weniger konkrete Aussagen getroffen werden, weshalb zu ihnen zu Beginn eines Gestaltungsprozesses Annahmen

getroffen werden, die es gilt, im Verlauf des Prozesses zu evaluieren, näher zu ergründen und konstant zu berücksichtigen. Antizipiert und zugeschrieben werden diese Eigenschaften vornehmlich vor dem Hintergrund der Notwendigkeit, Fortschritt im Gestaltungs-schehen zu ermöglichen. Das von mir verwendete Rollenkonzeptverständnis habe ich skizzenhaft in einem Entity-Relationship-Modell dargestellt (siehe Abb. 41).

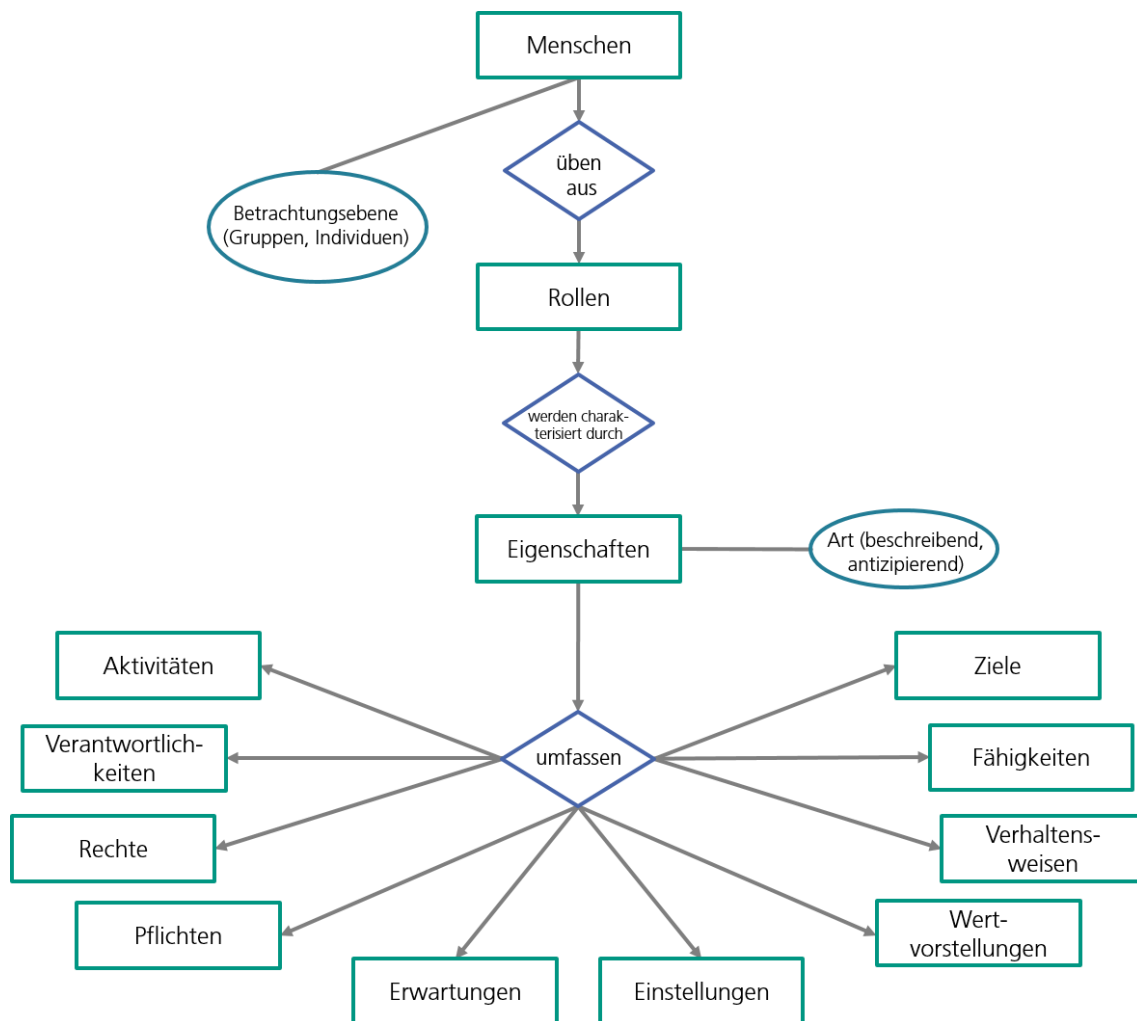


Abbildung 41: Entity-Relationship-Modell des von mir verwendeten Rollenverständnisses

## 5.3 Vorgehensweise

Für die Untersuchung der Rollenkonzepte in CTA und VSD führe ich eine vergleichende Analyse einschlägiger Literatur durch. Die vergleichende Analyse der Dissertation schließt an den bisherigen vergleichenden Diskurs beider Ansätze an (Doorn et al., 2013; Fisher et al., 2015; Grunwald, 2015; Pesch, 2015; Simon, 2017; Steen & Nauta, 2020; Wiarda, 2023, S. 39–58), in dem vergleichende Analysen dazu verwendet werden, um beispielsweise die Ziele der beiden Ansätze zu vergleichen, oder um zu untersuchen, inwieweit die Methoden der beiden Ansätze dazu befähigen, soziale und ethische Aspekte gestalterischen Handelns zu erkennen und über sie zu reflektieren. Im Unterschied zu diesen vergleichenden Analysen fokussiert die Dissertation die Rollenkonzepte in CTA und VSD, die nach meinem derzeitigen Kenntnisstand bislang noch in keiner Untersuchung im Vordergrund stehen. Die vergleichende Analyse dient dem Nachvollziehen der in CTA und VSD vorhandenen Rollenkonzepte, sowie der mit ihnen verbundenen und durch sie zum Ausdruck gebrachten Annahmen über die Eigenschaften und Tätigkeiten der an Gestaltungsprozessen beteiligten und der von ihnen betroffenen Menschen. Die vergleichende Analyse ermöglicht es, die Besonderheiten und Zusammenhänge der in CTA und VSD verwendeten Rollenkonzepte zu erfassen und dadurch nachzuvollziehen, wie die Rollenkonzepte zu der in den Ansätzen vertretenen Auffassung von Gestaltungsprozessen ineinandergreifen (sollen) und welchen Beitrag spezifische Charakteristika einzelner Rollenkonzepte möglicherweise dazu leisten. Die Eigenschaften und Funktionen der Rollenkonzepte sowie ihre Instrumentalisierung in CTA und VSD repräsentieren zudem, auf welche Weise beide Ansätze eine Menschorientierung interpretieren und wie diese unter Verwendung der Rollenkonzepte in Gestaltungsprozessen zu realisieren gesucht wird. In Summe liefert die Betrachtung der Rollenkonzepte wertvolle Erkenntnisse, mit deren Hilfe es mir möglich wird, fundierte Aussagen über die konzeptionelle Anschlussfähigkeit beider Ansätze zur Debatte um die MZKG in HCAI treffen zu können.

### 5.3.1 Durchführung

Als Datengrundlage dient der vergleichenden Analyse Literatur, die aus erster Hand derjenigen Personen stammt, die maßgeblich an der Ausarbeitung und Weiterentwicklung beider Ansätze beteiligt waren. Für CTA handelt es sich hierbei vornehmlich um Texte unter (Co-)Autorenschaft von Arie Rip, Johan Schot, Douglas Robinson und Haico te Kulve, für VSD zählen hierzu vornehmlich Texte unter (Co-)Autorenschaft von Batya Friedman, Peter Kahn und David Hendry sowie Lisa Nathan und Janet Davis. Die Orientierung an diesen Texten erlaubt es mir, die Bedeutung und Verwendung der Rollenkonzepte in CTA und VSD möglichst anhand ihrer originären Bedeutungen zu erfassen, sowie die mit ihnen verbundenen Intentionen, Gedankengänge und Argumentationslinien „im Sinne des Erfinders“ nachzuzeichnen. Zusätzlich zu diesem Literaturkorpus ist für die vergleichende Analyse Literatur von Interesse, die eine mehr reflektierende Position zu beiden Ansätzen einnimmt, und aufgrund einer größeren Distanz zu den Ursprungsumfeldern zusätzlich relevante Aspekte zu den Rollenkonzepten in den Vordergrund rückt.

Für die Durchführung der vergleichenden Analyse entwickle ich ein Kategoriensystem, welches ich in Tabellenform realisiere. Durch die Verwendung des Kategoriensystems wird es mir möglich, die vergleichende Analyse strukturiert durchzuführen, zentrale Aspekte der Rollenkonzepte zu erfassen und durch das Hinterlegen entsprechender Textpassagen im Kategoriensystem die Verwendungszusammenhänge der Rollenkonzepte im jeweiligen Ansatz nachzuzeichnen und nachzuvollziehen. Das Kategoriensystem umfasst mehrere Spalten, die verschiedene Funktionen erfüllen. Übergeordnete Vergleichskategorien ermöglichen die grundlegende Unterscheidung zwischen mehr theoretischen und mehr methodischen Elementen beider Ansätze, in deren Zusammenhang Rollenkonzepte von Bedeutung sind. Diese Vergleichskategorien werden stellenweise durch Subkategorien erweitert, um präzisere Analysen auf einer detaillierteren Ebene zu ermöglichen, oder um zu ermöglichen, Besonderheiten bestimmter Rollenkonzepte zu ergründen und hervorzuheben. Das Kategoriensystem unterscheidet mit zwei weiteren Spalten zudem zwischen den beiden Ansätzen CTA und VSD, um die Inhalte, die Rollenkonzepte betreffen, zu einem bestimmten Ansatz zuzuordnen. Jeweils eine weitere Spalte dient dazu, aussagekräftige Zitate aus der Literatur zu hinterlegen, um die Bedeutung eines die Rollenkonzepte betreffenden Gesichtspunkts zu erläutern, seine Relevanz für die jeweilige (Sub-)Kategorie hervorzuheben und um für die Durchführung der Analyse als Gedächtnisstütze zu dienen. Die nachfolgende Tabelle (siehe Tab. 2) zeigt das Layout des Kategoriensystems, zusammen mit zwei beispielhaften Einträgen, die von zwei Subkategorien Gebrauch machen.



Ver- gleichs- katego- rie	Subka- tegorie 1	Subka- tegorie 2	CTA	Quelle CTA	VSD	Quelle VSD
Zentrale Konzepte	Verwen- dete Rollen- konzep- te	Gestalter & Nutzer	„The simple contrast between technology developers and users is inapplicable then. Interaction and mutual learning become important to overcome mismatches and fragmentation, in innovation as well as in ‘distributed responsible development’. New ‘divisions of moral labour’ have to be invented, and one can see various actors exploring (even if reluctantly) possibilities.“	Rip & Robinson, 2013, S.42	“It recognizes designers as far more important professionals since they not only can provide us with technical means but can also address the values of people and society and think about expressing them in material culture and technology. Moreover, technological innovation is no longer a development that is separate from the values that users and society hold.“	van den Hoven et al. 2015, S.2-3
Vorge- schlagene Metho- den	Work- shops	Strategy articulation workshops	„and (2) the organizing and orchestration of workshops with a broad variety of stakeholders. The scenarios help to structure the discussion in the workshops (Robinson 2010) and stimulate learning about possible strategies.“	Rip & Robinson 2013, S.38		
		Prototyping-Workshops			„Development, analysis, and co-design of mock-ups, prototypes, and field deployments to scaffold the investigation of value implications of yet-to-be-built or widely adopted technologies. Mock-ups, prototypes, or field deployments emphasize implications for direct and indirect stakeholders, value tensions, and technology situated in human contexts.“	Friedman & Hendry, 2019, S. 59

Tabelle 2: Layout des Kategoriensystems der vergleichenden Analyse mit Beispieleinträgen

Das Kategoriensystem wird im Verlauf der vergleichenden Analyse sukzessive erweitert und mit entsprechenden Einträgen und Zitatstellen befüllt. So entsteht ein umfangreicher Korpus von Vergleichs- und Subkategorien, der die Eigenschaften und Besonderheiten der verschiedenen Rollenkonzepte wiedergibt und darüber hinaus Einblicke in die übergeordneten Denkstrukturen und Arbeitsweisen in beiden Ansätzen ermöglicht. Exemplarische Vergleichs- und Subkategorien fasst die nachstehende Tabelle (siehe Tab. 3) zusammen. Darin gebe ich auch Beispiele für Subkategorien wieder, sofern ich im Rahmen der vergleichenden Analyse Subkategorien entwickelt habe.

Vergleichskategorie	Subkategorie 1 (Beispiele)	Subkategorie 2 (Beispiele)
Verwendete Rollenkonzepte	Gebräuchlich in Gestaltungskontexten	Developers, Designers, Engineers, Users, Stakeholders
	Charakteristisch für CTA/ VSD	Actors, CTA agents, indirect stakeholders, impactees, comparative selectors
Attribute der Rollenkonzepte	Actor, Developer, Impactee	Relevanz, technische Fähigkeiten, Platzhalter, Bindeglied
Besonderheiten von Rollenkonzepten	CTA agent, comparative selector	Befugnisse, Aufgaben, Skillsets
	Indirect stakeholders, Impactees	Notwendigkeit für Berücksichtigung und Partizipation
	„Dualität“	Direct und indirect stakeholders, impactors und impactees, enactors und selectors
Ursprünge der Rollenkonzepte	Organisationsforschung	Change Agents
Methodische Operationalisierung von Rollenkonzepten	CTA agents, Designer, Stakeholder	CTA Workshops, Technical investigation
(Verbleibende) Offene Fragen	Festlegung von Grenzen/ Gültigkeitsbereichen	Wessen Interessen und Werte werden aufgegriffen?
	Lösungsorientierung	Umgang mit value conflicts

Tabelle 3: Zentrale Kategorien der vergleichenden Analyse im Hinblick auf Rollenkonzepte

### 5.3.2 Auswertung

Die Darstellung der Ergebnisse der vergleichenden Analyse orientiert sich an den ersten drei Forschungsfragen der Dissertation. In einem ersten Schritt werde ich die in CTA und VSD verwendeten Rollenkonzepte und die mit ihnen intendierten Zwecke beschreiben (Forschungsfrage 1). Dies etabliert ein Grundverständnis über die verschiedenen Rollenkonzepte, zeigt erste Verbindungslinien und Unterschiede zwischen den Rollenkonzepten sowie der Zugänge beider Ansätze zu Gestaltungsprozessen auf, und vermittelt die Charakteristika und Besonderheiten der verwendeten Rollenkonzepte. Auch werden hier bereits erste Hinweise zu Anknüpfungspunkten zum HCAI-Diskurs erkennbar, da die Bedeutungen der Rollenkonzepte Aufschluss über die Art der in CTA und VSD angestrebten Menschorientierung geben. Anschließend werde ich auf die Ursprünge der verwendeten Rollenkonzepte eingehen (Forschungsfrage 2), die vertiefende Erkenntnisse zu den

Verwendungsweisen und den Verwendungsgründen in den Ansätzen bereitstellen. Die Betrachtung der Ursprünge rückt Zusatzinformationen in den Vordergrund, die bei einer reinen Beschreibung der Rollenkonzepte und ihrer Verwendungszwecke leicht übersehen und daher möglicherweise verborgen bleiben würden. So wird das bereits erarbeitete Verständnis der Rollenkonzepte durch punktuelle, inhaltliche Vertiefungen erweitert und komplementiert. Das damit erarbeitete Wissen über die Rollenkonzepte ermöglicht mir die Identifikation blinder Flecke der Rollenkonzepte (Forschungsfrage 3). Die blinden Flecke geben Aufschluss darüber, an welchen Stellen die Ansätze und die in ihnen verwendeten Rollenkonzepte mit Blick auf den MZKG-Diskurs Defizite aufweisen. Für die Fragestellung der Dissertation ist es bedeutsam zu ergründen, wie sich diese blinden Flecke äußern und wie sie im Hinblick auf menschenzentrierte Gestaltungsprozesse zu bewerten sind. Die Betrachtung der blinden Flecke liefert somit weitere entscheidende Hinweise, um die Anschlussfähigkeit beider Ansätze zu Fragen und Herausforderungen in HCAI beurteilen zu können.

## 5.4 Verwendete Rollenkonzepte

In den beiden Ansätzen CTA und VSD werden verschiedene Rollenkonzepte verwendet, um die an Gestaltungsprozessen beteiligten bzw. von diesen betroffenen Menschengruppen zu beschreiben. Die verschiedenen Bedeutungen und intendierten Funktionen der Rollenkonzepte lege ich nachfolgend dar, wobei sich die Struktur der Darlegung am Spezifikitätsgrad der Rollenkonzeptbeschreibungen in der Literatur orientiert. Zu Beginn erläutere ich die mehr übergreifenden Rollenkonzepte, an die sich die Darstellung der detaillierteren Rollenkonzepte anschließt. Letztere können prinzipiell in beiden Ansätzen Präzisierungen der übergeordneten Rollenkonzepte darstellen, die in den Ansätzen aus diversen Gründen vorgenommen werden. Manche der Rollenkonzepte finden in beiden Ansätzen Verwendung, manche Rollenkonzepte sind jedoch charakteristisch für lediglich einen der beiden Ansätze. In Summe werden durch die Rollenkonzepte die verschiedenen Perspektiven deutlich, die in beiden Ansätzen hinsichtlich menschlicher Aktivitäten in Gestaltungsprozessen eingenommen und als bedeutsam erachtet werden. Konkret wird in den beiden Ansätzen dabei zwischen (1) den Involvierten, (2) den informierten Einflussnehmern, (3) den Unbeteiligten und Betroffenen, (4) den Gestaltern, (5) den intervenierenden Prozessbegleitern, (6) den Nutzern, (7) den Technikenthusiasten und (8) den Technikgegnern unterschieden (siehe Tab. 4).

Rollenkonzeptgruppe	Bezeichnete Rollenkonzepte
Die Involvierten	Stakeholders, Actors
Die informierten Einflussnehmer	Impactors, Insider, Enactors, comparative Selectors
Die Unbeteiligten und Betroffenen	Impactees, Indirect Stakeholders
Die Gestalter	Developers, Engineers, Designer
Die intervenierenden Prozessbegleiter	CTA agents
Die Nutzer	Users
Die Technikenthusiasten	Promoters, Proponents
Die Technikgegner	Opponents

Tabelle 4: Übersicht der verwendeten Rollenkonzepte in CTA und VSD

### 5.4.1 Die Involvierten

Die Involvierten bezeichnet eine Gruppe von Rollenkonzepten, die Menschen als am Gestaltungsprozess beteiligt konzipieren, die Art oder den Grund der Beteiligung häufig jedoch nicht näher bestimmen. Bei den Involvierten handelt es sich somit um die im Vergleich mit anderen Rollenkonzepten am stärksten variablen und am stärksten impliziten Rollenkonzepte, da sich ihre konkrete Bedeutung häufig nur im Zusammenhang des Verwendungskontexts erschließen lässt. Zu den Rollenkonzepten der Involvierten zählen die Stakeholder und die Akteure, die beide ähnliche Eigenschaften besitzen.

#### Stakeholder

In CTA wie auch in VSD wird das Rollenkonzept der Stakeholder verwendet. Stakeholder besitzen Relevanz für einen Gestaltungsprozess, die Gründe für diese Relevanz können jedoch stark variieren. So können Stakeholder beispielsweise als relevant erachtet werden, da sie ein hohes wirtschaftliches oder politisches Interesse am Fortschritt oder Ausgang eines Gestaltungsprozesses besitzen, oder da das Wissen über das sie verfügen für den Gestaltungsprozess als wertvoll oder erforderlich erachtet wird. Auch kann sich die Bedeutung von Stakeholdern aus der chronologischen Betrachtung eines Gestaltungsprozesses erschließen. So ist es beispielsweise möglich, dass Stakeholder die Impulsgeber für einen Gestaltungsprozess darstellen, oder dass sie während bestimmter Phasen oder nach Abschluss eines Gestaltungsprozesses Nutznießer oder Betroffene der durch ein technisches System zur Verfügung gestellten Informationen oder Handlungsspielräume darstel-

len. Da die Art der Beteiligung sowie die Beteiligungsintensität von Stakeholdern stark variieren kann, wird ihre konkrete Bedeutung an vielen Stellen der Literatur lediglich angedeutet. Eine vergleichsweise konkrete Beschreibung des Rollenkonzepts in CTA findet sich bei Franks und Cohen, die Stakeholder beschreiben als „those affected by, or that can effect, the technology, operation or event“ (Franks & Cohen, 2012, S. 1231). In VSD wird der Auseinandersetzung mit dem Rollenkonzept mehr Aufmerksamkeit gewidmet, wie beispielsweise bei Friedman und Hendry deutlich wird:

*„Defining stakeholders in terms of roles positions designers to understand how the same “entity” could hold multiple relationships to the same technology, as well as the potentially dynamic nature of an entity’s relationship to a particular technology“* (B. Friedman & Hendry, 2019, S. 40).

Diese Verwendung des Rollenkonzepts steht somit sinnbildlich für die Bedeutung der interactional stance in VSD. Die Anerkennung und der Umgang mit den vielfältigen, dynamischen Interaktionsbeziehungen zwischen Mensch und Technik dient zudem der in VSD bedeutsamen Unterscheidung der direkten und indirekten Stakeholder. Unter direkten Stakeholdern werden dabei Gruppen und Individuen verstanden, die auf direkte Weise mit einem technischen System interagieren und dadurch Zustandsveränderungen bewirken, die Auswirkungen auf ihre eigenen Leben sowie auf die der indirekten Stakeholder haben. Indirekte Stakeholder bezeichnen sämtliche weitere Gruppen und Individuen, die nicht dazu in der Lage sind, mit einem System zu interagieren und sich deshalb häufig in der Position befinden, sich mit den durch die direkten Stakeholder hervorgerufenen Auswirkungen und geschaffenen Zustände abzufinden und sich an diese anzupassen. Aufgrund ihrer mehr mittelbaren Beziehung zum System und deshalb vermeintlich mittelbaren Bedeutung im Gestaltungsprozess werden sie häufig von Gestaltungsverantwortlichen ignoriert (B. Friedman, Kahn, & Borning, 2006, S. 361). Mit dem Hinweis auf das Ignorieren indirekter Stakeholder äußert VSD eine grundsätzliche Kritik anderer Gestaltungsansätze, die es durch die bewusste Unterscheidung der indirekten und direkten Stakeholder zu adressieren sucht (siehe Kap. 5.4.3). In der VSD-Literatur finden sich verschiedene Beispiele um die Beziehung zwischen direkten und indirekten Stakeholdern zu illustrieren. Bei Friedman et al. (2006) werden beispielsweise im Kontext der Gestaltung eines Systems zur Verwaltung von Patientenakten die Patienten selbst und ihre Angehörigen als indirekte Stakeholder beschrieben, während verschiedene Berufsbilder des Krankenhauspersonals sowie Krankenversicherungen als direkte Stakeholder eines Systems identifiziert werden, nach deren Bedürfnissen und Werten ein System typischerweise gestaltet wird.

Abgesehen von der bewusst durchgeführten Unterscheidung direkter und indirekter Stakeholder in VSD wird das Rollenkonzept in beiden Ansätzen dennoch überwiegend vage verwendet. Jedoch liegt genau in dieser Vagheit die intendierte Funktion des Rollenkonzepts. Es dient in CTA und VSD primär dazu, sämtliche für den Gestaltungsprozess als relevant erachteten Gruppen und Individuen zu bezeichnen und ähnlich einer Art größtem gemeinsamen Nenner zu repräsentieren. Bei der Verwendung des Rollenkonzepts steht in der Regel weniger im Vordergrund, auf präzise Weise Rollenmerkmale zu

spezifizieren oder zu adressieren, sondern es wird mehrheitlich dazu verwendet, um mit ihm in Anspruch zu nehmen, eine maximal vollständige Anzahl relevanter Gruppen und Individuen identifiziert zu haben, und diese mit der Verwendung des Rollenkonzepts zu bezeichnen und zu adressieren. Somit geht mit dem Rollenkonzept häufig ein implizierter Vollständigkeitsanspruch einher, der sich jedoch lediglich auf einen bestimmten Kontext, Gestaltungsprozess oder Teilprozess bezieht. Das Rollenkonzept besitzt auch eine vereinende und inkludierende Funktion, um sämtliche relevante Gruppen und Individuen einzuschließen. Diese Funktion erfährt das Rollenkonzept durch das Minimum an Attributen, die typischerweise verwendet werden, um Eigenschaften von Stakeholdern zu klassifizieren. Ähnlich wie die grundsätzliche Bedeutung des Rollenkonzepts werden seine Attribute in der Literatur nur an wenigen Stellen expliziert. Ein Beispiel findet sich in der VSD-Literatur, wo besonders zwei Attribute des Rollenkonzepts hervorgehoben werden: „We call attention to two characteristics of stakeholders: multiple roles and fluctuating boundaries“ (Nathan et al., 2008, S. 3). Der Hinweis auf die Attribute der Rollenpluralität und der fließenden Bedeutungsgrenzen erfolgt in VSD besonders vor dem Hintergrund, um auf die verschiedenen Interaktionsbeziehungen zwischen Mensch und Technik hinzuweisen, diese untersuchen zu können, sowie um die zentrale Unterscheidung der direkten und indirekten Stakeholder durchzuführen. Solch eine häufig geringe Anzahl an Attributen verleiht dem Rollenkonzept seine hohe Flexibilität, da die wenigen Attribute es erlauben, sich mithilfe weniger aber dennoch als gültig angenommener Kriterien der in Gestaltungsprozessen typischerweise vorhandenen Rollenvielfalt und sozialen Konstellationen anzunähern, diese im Verlauf des Prozesses zu berücksichtigen und somit insgesamt ihrer hohen Dynamik zu begegnen. Das Rollenkonzept erlaubt es zudem die wenigen Attribute bei Bedarf zu erweitern, näher zu spezifizieren oder gänzlich durch andere Attribute zu ersetzen, sollte ein bestimmter Umstand im Gestaltungsprozess oder die Darstellung eines bestimmten Gedankengangs dies erfordern. Somit ist das Rollenkonzept der Stakeholder auch ein zentrales Instrument für beide Ansätze, um ihrem Anspruch gerecht zu werden, zur Gestaltung sämtlicher Technologien in sämtlichen sozialen Konstellationen und Situationen angewandt werden zu können.

Mit den Stakeholdern werden in CTA und VSD ein nach Ansicht innerhalb der Ansätze adäquates und maximal flexibles Rollenkonzept verwendet, das es ermöglicht, sämtliche relevante Gruppen und Individuen zu identifizieren, zu adressieren und bedarfsweise genaueren Betrachtungen zu unterziehen. Diese Flexibilität kann vor allem in Gestaltungsprozessen Bedeutung erlangen, in denen zu Beginn nur wenig Informationen über bestimmte Gruppen oder Individuen zur Verfügung stehen und zu denen deshalb Annahmen getroffen werden müssen, um grundlegende Gestaltungsentscheidungen treffen zu können. Durch die (zeitweise) Verwendung des Stakeholder-Rollenkonzepts können im weiteren Verlauf tatsächliche, konkrete Attribute jener Gruppen und Individuen identifiziert werden, die die zuvor lediglich angenommenen Attribute ergänzen oder ersetzen („Platzhalterfunktion“). Stakeholder können als abstrakte Platzhalter fungieren, deren konkrete Ausprägungen sich durch die in einem Gestaltungskontext vorgefundenen Zusammenhänge ableiten und/ oder für die dort notwendigen Gegebenheiten angepasst werden können. In Verbindung mit dem angenommenen grundlegenden Gültigkeitsanspruch gewährleistet das Rollenkonzept so trotz gewisser Unsicherheiten den Fortgang

des Gestaltungsprozesses. In dieser Interpretation verwendet, wird in VSD dadurch auch dem mittleren Namensbestandteil „sensitive“ gerecht, da mithilfe des Rollenkonzepts die Kontextsensitivität von Gestaltungsprozessen anerkannt werden kann, um sie produktiv in Gestaltungsprozesse zu integrieren.

Nachvollziehbar wird die Praktikabilität der Offenheit und Flexibilität des Rollenkonzepts auch beim Blick auf die Methoden, die in beiden Ansätzen vorgeschlagen und verwendet werden. In CTA sieht beispielsweise die Durchführung der „Multi-Stakeholder Workshops“ (Robinson, 2009, S. 1225) vor, Stakeholder aus verschiedenen für einen Gestaltungsprozess relevanten Gesellschafts- oder Professionsbereichen zusammenzuführen, um unter den Stakeholdern ein „learning about other perspectives“ (Robinson, 2009, S. 1237) zu fördern und sie bei der Diskussion soziotechnischer Szenarien zu unterstützen. Um dieses Ziel zu erreichen, wird eine möglichst hohe Vielfalt an Stakeholderperspektiven benötigt, da nur ein gewisser Grad an Perspektivenpluralität die angestrebten kontroversen und deshalb fruchtbaren Diskussionen verspricht. Die Verwendung des Stakeholder-Rollenkonzepts zur Beschreibung der Methode trägt dazu bei, dass auch wenig mit CTA vertraute Anwender eine möglichst hohe gedankliche Offenheit bei der Planung eines Workshops bewahren und die Identifikation und Rekrutierung möglicher Teilnehmer mit einem offenen Mindset betreiben. Die Offenheit des Rollenkonzepts ermöglicht es, die Relevanz möglicher Teilnehmer anhand vieler unterschiedlicher Kriterien auszumachen, wodurch die Identifikation und Auswahl möglicher Workshopteilnehmer unterstützt und zusätzlich die Wahrscheinlichkeit erhöht wird, eine zielführend hohe Anzahl an Teilnehmern zu rekrutieren.

In Ergänzung zu den dargelegten Eigenschaften besitzt das Rollenkonzept eine weitere Besonderheit in seiner Verwendung in VSD. Hier kann sich seine Bedeutung neben Menschen auch auf nicht-menschliche Entitäten beziehen:

*„Within value sensitive design, stakeholders are broadly conceived: they can be people, groups, neighborhoods, communities, organizations, institutions, or societies, and can also include past and future generations, nonhuman species, and other elements such as historic buildings or sacred mountaintops“ (B. Friedman & Hendry, 2019, S. 40).*

Die Ausdehnung auf nicht-menschliche Entitäten erweitert die zuvor dargelegte Flexibilität des Rollenkonzepts deutlich und besitzt Ähnlichkeiten zur Konzipierung relevanter Prozesselemente in der MSEP (siehe Kap. 5.2.2).

Eine Begründung für diese weite Konzipierung von Stakeholdern liefern Friedman und Hendry im direkten Nachfolgesatz, wo sie beschreiben, dass in VSD Gestaltungsverantwortliche dazu aufgefordert werden sollen, ein robustes Set an Stakeholdern zu erarbeiten, um mit dessen Hilfe diejenigen Stakeholder zu identifizieren, die in hohem Maße von der Gestaltung eines technischen Systems betroffen sind. Dies verdeutlicht den zuvor als „Platzhalterfunktion“ beschriebenen, zusätzlichen Zweck des Rollenkonzepts, der initial eine Annäherung an den Gestaltungskontext mithilfe weniger als angemessen angenommenen Attributen ermöglicht, um diese im weiteren Verlauf des Prozesses zu präzisieren. Trotzdem, dass hier die Möglichkeit dieser Interpretation und Verwendung auf nicht-menschliche Entitäten aufgezeigt wird, bezieht sich das Rollenkonzept in der VSD-Literatur primär auf Menschen.

### Actors

Das Rollenkonzept der Akteure weißt vergleichbare Eigenschaften und Verwendungszwecke wie das Rollenkonzept der Stakeholder auf und ist besonders in CTA stark präsent. Tatsächlich werden beide Rollenkonzepte in der CTA-Literatur stellenweise synonym verwendet, wie beispielsweise bei Kornelia Konrad, die erläutert, dass

*„viele CTA-Arbeiten einen Stakeholder-Ansatz verfolgen, d.h. Akteursgruppen gemäß ihrer sozialen Rolle unterscheiden“ (Konrad, 2021, S. 211).*

Mit dem Rollenkonzept der Akteure wird die Annahme zum Ausdruck gebracht, dass bei der Gestaltung technischer Systeme Menschen in verschiedenen Rollen oder Positionen handelnd aktiv sind. Kennzeichnend für Akteure sind in CTA ihre konkreten Ziele, Interessen und Werte (Schot & Rip, 1997, S. 257), die es im jeweiligen Gestaltungszusammenhang zu ermitteln gilt, um über sie zu reflektieren. Zusätzlich besitzen Akteure bestimmte Perspektiven auf die bestehenden und möglichen Entwicklungspfade einer Technologie. Zu diesen Perspektiven besteht in CTA die Annahme, dass Akteure primär ihre eigenen Perspektiven vertreten und sich aus einer intrinsischen Motivation kaum mit den Perspektiven anderer Akteure eines Technologiefelds auseinandersetzen. Diesen Zustand strebt CTA an, insbesondere durch Workshops zu adressieren um durch die Förderung eines Bewusstseins zur Perspektivenpluralität unter den Akteuren in einem Technologiefeld zu einer Gestaltung sozialverträglicher Technik beizutragen. Anhand dieses Attributminimums erschließt sich, weshalb Akteure für einen Gestaltungsprozess als relevant erachtet werden.

Bereits an dieser Charakterisierung des Rollenkonzepts werden die Parallelen zum Rollenkonzept der Stakeholder deutlich. Ähnlich wie bei Stakeholdern dient das Rollenkonzept der Akteure dazu, die in Gestaltungsprozessen gegebene Rollen- und Perspektivenvielfalt anzuerkennen und die damit verbundenen Dynamiken und Wechselwirkungen zu beschreiben und zu untersuchen. Während sich das Stakeholderverständnis in VSD stark auf ein Interaktionsverständnis der Mensch-Technik-Interaktion bezieht, bezieht sich das Akteursverständnis in CTA stark auf die in der Soziologie übliche Betrachtung sozialen Handelns durch den Akteursbegriff. Durch den Akteursbegriff wird in CTA häufig lediglich



zum Ausdruck gebracht, dass Menschen in dem sozialen Kontext eines Gestaltungsprozesses Handlungen durchführen, ohne die diesen Handlungen zugrundeliegenden Motivationen oder die damit verbundenen Ziele zu konkret zu beleuchten. Zu den Motivationen und Zielen von Akteuren werden in der Regel wenig konkrete Annahmen oder Aussagen getroffen, da über sie (durch eine Technologiefeldanalyse) häufig erst nähere Informationen ermittelt werden müssen. Entsprechend vage und überwiegend unspezifisch wird das Rollenkonzept in weiten Teilen der CTA-Literatur verwendet. Doch wie bei den Stakeholdern liegt auch hier darin der Nutzen für CTA, denn so kann eine Annäherung an und Begleitung von Gestaltungsprozessen sämtlicher Technologien erfolgen, bei der sämtliche für den Gestaltungsprozess als relevant erachteten Gruppen und Individuen bezeichnet werden (Vollständigkeitsanspruch). Keine möglicherweise relevanten Gruppen und Individuen werden ausgeschlossen (vereinende Funktion) und durch die nähere Auseinandersetzung mit Akteuren im Verlauf eines Gestaltungsprozesses können die allgemeinen Attribute der Ziele, Interessen, Perspektiven usw. sukzessive durch konkretere und kontextspezifische Merkmale präzisiert werden („Platzhalterfunktion“). Das Rollenkonzept der Akteure impliziert, sämtliche für den Gestaltungsprozess relevanten Gruppen und Individuen zu bezeichnen, ohne diese näher auf bestimmte Merkmale einzugrenzen und es bietet ausreichend Spielraum um die für einen Gestaltungskontext spezifischen Merkmale handelnder Akteure zu erfassen (hohe Flexibilität).

Auch wenn im Großteil der CTA-Literatur das Rollenkonzept implizit verwendet wird, finden sich punktuell detaillierte Beschreibungen. Johan Schot (1992, S. 42) unterscheidet beispielsweise drei Arten von Akteuren: (1) diejenigen, die Ziele und Heuristiken für die Gestaltungsvorgänge formulieren, (2) diejenigen, die von Ausschnitten sozialer Realität außerhalb des Gestaltungsprozesses diesen zu beeinflussen suchen, sowie (3) diejenigen, die sowohl an der Erarbeitung wie auch Auswahl der möglichen Gestaltungspfade beteiligt sind. Auf einer höheren Abstraktionsebene unterscheiden Schot und Rip (Schot & Rip, 1997, S. 256) *Technology actors*, **societal actors** und **„actors at a meta level“**. Unter *Technology actors* verstehen sie Unternehmen, Behörden oder Labore, die in TG investieren oder diese ausführen. *Societal actors* zeichnen sich dadurch aus, dass sie technologische Entwicklungen antizipieren und versuchen diese durch beispielsweise Regulierungsmaßnahmen oder Bildungsprogramme zu beeinflussen. Entsprechend können je nach Gestaltungskontext Regierungsbehörden, Unternehmen oder andere soziale Gruppen wie beispielsweise Vereine oder Stiftungen Konkretisierungen sozialer Akteure darstellen. Als **„actors on a meta level“** werden Gruppen und Individuen verstanden, die in der Lage sind, zu bestimmten Zeitpunkten im Gestaltungsprozess Urteile und Entscheidungen zu fällen und die über die zuvor dargelegten Unterschiede der *Technology actors* und *societal actors* im Bilde sind. Regierungen sind hierfür das am häufigsten genannte Beispiel in der CTA-Literatur, jedoch können auch TA-Einrichtungen Konkretisierungen solcher Meta-Akteure darstellen.

Diese „Dreiteilung“ der an Gestaltungsprozessen beteiligten Rollen in TG (*technology actors*), Gesellschaft (*societal actors*) und Regulierung (*actors on a meta level*) repräsentiert eine fundamentale, tiefreichende Annahme in CTA und zieht sich entsprechend durch den gesamten CTA-Literaturkorpus. Die Unterscheidung dieser drei Akteursgruppen hängt

zusammen mit einem weiteren in CTA zentralen Prinzip, was als „managing technology in society“ (Rip et al., 1995). bezeichnet wird. Damit wird in CTA untersucht und beschrieben, auf welche Institutionen und Rollen in Gesellschaften sich die Verantwortung für die mit der Gestaltung und Aneignung von Technik verbundenen Folgen verteilen lässt. Die CTA-Position zu dieser kniffligen Fragestellung besteht darin, dass sich diese Verantwortung nicht ausschließlich der rechtsprechenden Regulierungsbehörden von Regierungen oder den technikgestaltenden Unternehmen zuordnen lässt, sondern dass sich diese Verantwortung auf alle drei Akteursgruppen verteilt:

*„This implies not only that responsibility for “managing” technology in society is distributed over more actors, but also that the nature of the responsibility is different for different kinds of actors. In a sense, all three kinds of actors are necessary“* (Schot & Rip, 1997, S. 257).

Hieran wird die Kritik in CTA hinsichtlich des two-track-approachs (siehe Kap. 4.2.1) erkennbar, da darin die Gesellschaft respektive die sozialen Akteure kaum Berücksichtigung finden. Ihrer Betrachtung wird in CTA eine hohe Aufmerksamkeit geschenkt.

In der VSD-Literatur ist das Rollenkonzept der Akteure deutlich weniger stark vertreten, es findet jedoch auch hier punktuell Anwendung, um die verschiedenen beteiligten Rollen und die daraus entstehenden Rollenverhältnisse in Gestaltungskontexten zu beschreiben. So wird es beispielsweise dazu benutzt, um die unterschiedlichen Rollen technikgestaltender und technikaneignender Gruppen und Individuen zu beschreiben:

*„clients, co-designers, designers, field-testers, inventors, and so forth imagine and develop technology that, later, is appropriated and integrated into society by other kinds of actors“* (B. Friedman & Hendry, 2019, S. 28–29).

Hier wird eine kategorische Unterscheidung technikgestaltender und technikaneignender Akteursgruppen vorgenommen, um im Zusammenhang mit den in VSD vorgeschlagenen Methoden die gestaltenden Akteure in die Lage zu versetzen, sich die aneignenden Akteure zu vergegenwärtigen (envision). Verbunden damit ist der Hinweis, dass sich deren Beziehungen zum technischen System im Laufe eines Gestaltungsprozesses verschiedentlich entfalten und verändern können. Das Envisioning ist dabei ein Begriff aus dem Bereich des Projektmanagements (Vorbereitungsphase), das in VSD stellenweise aufgegriffen wird, um zu verdeutlichen, welcher Mehrwert eine möglichst ganzheitliche Sichtweise auf Gestaltungsprozesse beinhaltet:

*„Rarely does a single actor, design team, or even a governmental agency, have access to a whole view of the system, and it is even more unlikely that they have control over all of the parts“* (Nathan et al., 2008, S. 2).

An diesen beiden Beispielen wird wieder die hohe Abstraktionsebene des Rollenkonzepts sowie seine Verwendung zu mehr sozialwissenschaftlichen Betrachtungen deutlich, wovon letztere in weiten Teilen der VSD-Literatur dann durchgeführt werden, um konkrete Aspekte zur Sprache zu bringen. Ein Beispiel hierfür findet sich bei Friedman und Hendry, die

mit dem Rollenkonzept arbeiten um VSD-Positionen mit Gedanken zu Resilienz zu verknüpfen:



*„When applying the resilience grammar, some set of actors will need to be identified to make these decisions. That actor or set of actors could be policymakers, community members, technologists, or others, potentially in combination“* (Hendry & Friedman, 2023, S. 8).

Abgesehen vom Stakeholderrollenkonzept wird in VSD im Vergleich mit CTA deutlich mehr versucht, die verwendeten Rollenkonzepte zu präzisieren und einen Anwendungsbezug herzustellen. Dies wird beispielsweise bei Jacobs et al. (2021) deutlich, die in ihrer Arbeit das Rollenkonzept der Akteure dazu verwenden, um mit seiner Hilfe die Machtverteilung in Gestaltungsprozessen von zunehmend komplexen soziotechnischen Ökosystemen zu untersuchen. Ihnen dient das Rollenkonzept dazu, verschiedene relevante Akteursgruppen für ihre Studie zu identifizieren und zu unterscheiden. Ein weiteres Beispiel stellt die Arbeit von van Wynsberghe (2013) dar, wo das Rollenkonzept ebenfalls dazu verwendet wird, um zwischen den verschiedenen involvierten Rollen eines Gestaltungsprozesses zu unterscheiden. Der Autorin dient das Rollenkonzept dazu, um die den von Ingenieuren getroffenen Gestaltungsentscheidungen innewohnende Verantwortung hervorzuheben sowie um zwischen menschlichen und nicht-menschlichen Akteuren zu unterscheiden:

*„By making choices about what should and should not be delegated to certain actors (human or nonhuman), engineers may change the distribution of responsibilities in a network“* (van Wynsberghe, 2013, S. 412–413).

Hier wird durch die Unterscheidung menschlicher und nichtmenschlicher Akteure deutlich, dass in VSD das Rollenkonzept der Akteure nur mit geringfügigen Unterschieden verwendet wird. Dieser Eindruck wird auch dadurch bestätigt, dass Friedman und Hendry selbst den Akteursbegriff und den Stakeholderbegriff weitestgehend synonym verwenden.

#### 5.4.2 Die informierten Einflussnehmer

Die informierten Einflussnehmer wie auch sämtliche nachfolgende Rollenkonzeptgruppen stellen in vielen Verwendungskontexten der CTA- und VSD-Literatur bewusst vorgenommene Präzisierungen der Rollenkonzepte der Stakeholder oder Akteure dar. Zu den informierten Einflussnehmern zählen dabei die Rollenkonzepte der Impactors, der Insider und der Encactors, die ausschließlich in CTA verwendet werden und in VSD keine Verwendung finden. Sie dienen in CTA vornehmlich der Gegenüberstellung zu den Rollenkonzepten der Impactees, den Outsidern und den comparative Selectors, um die Beziehungen und Unterschiede zwischen diesen Rollenkonzepten hervorzuheben und um die sie auszeichnenden, häufig invers gelagerten Interessen und Intentionen darzustellen. Die informierten Einflussnehmer verfügen über spezifisches Wissen, das ihren Gegenstück-Rollenkonzepten nicht zur Verfügung steht und das sie häufig dazu verwenden, um den Fortgang und die Resultate eines Gestaltungsprozesses zugunsten ihrer Ansichten

und Ziele zu beeinflussen. Impactors, Insiders und Encators werden in der CTA-Literatur stellenweise synonym verwendet, wodurch sich die sie auszeichnenden Eigenschaften überlagern. Bedeutsam wird ihre Unterscheidung dann, wenn in CTA gedankliche Perspektivwechsel durchgeführt werden, da Impactors ihre Einflussnahme stark durch das (direkte) Gestalten in Gestaltungsprozessen realisieren, während Enactors ihren Einfluss mehr durch regulierende und verordnende Maßnahmen realisieren. Die Unterscheidung der Rollenkonzepte dient damit auch der Hervorhebung unterschiedlicher Zeitpunkte eines Gestaltungsprozesses.

## Impactors

Das Rollenkonzept der Impactors zeichnet sich in CTA dadurch aus, dass es Gruppen und Individuen beschreibt, die direkt oder sehr nahe an den in einer Gesellschaft vorhandenen Stellen und Institutionen verortet sind, wo wegweisende technische Entwicklungen herbeigeführt werden. Impactors nehmen an solchen zentralen Weichenstellen Einfluss auf die dort stattfindende Gestaltung technischer Systeme, jedoch geht aus der CTA-Literatur nur selten hervor, wie sich diese Einflussnahme konkret vollzieht. Ein zentraler Zweck des Rollenkonzepts besteht darin, die in einer Gesellschaft vorhandenen Wissens- und Machtunterschiede hervorzuheben, die sich auch auf Gestaltungsprozesse technischer Systeme übertragen. In CTA-Projekten wird deshalb häufig angestrebt, die Orte und die Zeitpunkte der Gewinnung dieser das Ungleichgewicht auszeichnenden, zentralen Informationen zu ermitteln sowie die dortigen Informationsflüsse nachzuvollziehen. Im Gegensatz zu seinem Gegenstück-Rollenkonzept, den von technischen Entwicklungen betroffenen Impactees (siehe Kap. 5.4.3), verfügen Impactors immer über verschiedene Vorteile hinsichtlich Zeitkontingenten und Informationen, da sie früher und genauer über technische Entwicklungen und mögliche künftige Entwicklungspfade Bescheid wissen:

*„This can be a difference in power, but is always a difference in timing. Initiators of technological development know more, and have invested more, at an early stage, and impactees and spokespersons for society have to wait and in a sense, follow their lead“* (Rip, 2008, S. 146).

Stellenweise können Impactors diese technischen Entwicklungen auch selbst herbeigeführt haben, doch auch unabhängig von ihrem direkten Involvement in der Herbeiführung verfügen sie im Vergleich zu den Impactees immer über die Möglichkeit, einer sehr direkten Einflussnahme künftiger Entwicklungen ab einem bestimmten Zeitpunkt im Prozess. Durch ihren zeitlichen Vorteil können sie entsprechend ausgefeilte Argumentationsstrategien für die fortgesetzte Gestaltung nach ihren Vorstellungen ausarbeiten oder diese bereits in Maßnahmen überführen, wie beispielsweise die Anschlussfinanzierung bestimmter Anwendungen einer Technologie.

Eine vornehmliche Funktion des Rollenkonzepts der Impactors besteht somit darin, Macht- und Wissensunterschiede hervorzuheben, um davon ausgehend die Bedeutung der Ermittlung solcher Macht- und Wissenszentren und damit verbundener Strukturen in Gestal-

tungsprozessen darzulegen. Dazu wird häufig die Gegenüberstellung zu den Impactees bemüht, was auch bei den Rollenkonzepten der Insider und Enactors deutlich wird und an die grundlegende Unterscheidung erinnert, die in vielen Softwareentwicklungsprozessen zwischen Gestaltern und Nutzern vorgenommen wird. Die Bedeutung des einen Rollenkonzepts erschließt sich primär in der Betrachtung der Zusammenhänge und Unterschiede des gegenüberliegenden Rollenkonzepts, um die zwischen beiden bestehenden Grenzen und in CTA besonders der „Asymmetrien“ hinsichtlich Interessen, Möglichkeiten, Wissensbeständen usw. nachzuzeichnen. Auf eine weitere Funktion des Rollenkonzepts verweist Imre Hronszky (2005), indem er das Rollenkonzept in Bezug zu der Vorstellung gesellschaftlichen Wandels und Fortschritts als evolutionäre Wettläufe (races) setzt. Werden TG-Prozesse als solche Wettläufe betrachtet, kann nachgezeichnet werden, dass sie häufig Gewinner und Verlierer zum Ergebnis haben. Gewinner befeuern und gestalten den Verlauf nach ihren Vorstellungen, während sich die Verlierer überwiegend mit den negativen Effekten und Konsequenzen solcher Wettläufe konfrontiert sehen. Diese Unterscheidung von Gewinnern und Verlieren spiegelt sich in CTA bei den Impactors und Impactees wieder, „in what Rip calls, the relation of „impactors“ (those at the source of impacts) and „impactees““ (Hronszky, 2005, S. 50). Die mit dieser Analogie verbundene Interpretation gesellschaftlichen Wandels als evolutionäre Vorgänge lässt zudem erkennen, woher die in CTA vorhandene Interpretation von Gestaltungsprozessen als Koevolution von Wissenschaft, Technik und Gesellschaft stammt. Somit hat das Rollenkonzept der Impactors in CTA auch die Funktion, dieser (ko-)evolutionären Interpretation von Gestaltungsprozessen Rechnung zu tragen und durch seine Verwendung bestimmte CTA-typische Standpunkte und Argumentationslinien zu unterstützen. Zusätzlich wird hier auch wieder die Verbindung zu dem von Collingridge formulierten Kontrolldilemma im Zusammenhang der Gestaltung und Aneignung von Technik erkennbar. Wenn es in CTA gelingt, unter Zuhilfenahme des Rollenkonzepts der Impactors die Machtpositionen und Machtstrukturen in Gestaltungsprozessen zu identifizieren und transparent werden zu lassen, können die negativen Konsequenzen weiterer Entwicklungspfade aufgezeigt, abgemildert oder gänzlich abgewendet werden und somit die Anzahl möglicher Verlierer reduziert werden.

## Insider

Das Rollenkonzept der Insider dient in CTA dazu, sich besonders mit dem Wissens- und Informationsvorsprung auseinanderzusetzen, über den sie gegenüber Outsidern (siehe Kap. 5.4.3) verfügen. Dies dient in CTA dazu aufzeigen zu können, wie der Informationsvorsprung ihre erweiterte Machtposition etabliert:

*„Insiders who know much more about the technology and therefore position themselves as also more knowledgeable about its potential embedding in society“ (Rip, 2001, S. 111).*

Insider verfügen über Detailwissen zu Verfahren, Produktionsweisen, Abhängigkeiten usw. und wähnen sich deshalb häufig in der Position, entscheiden zu können, welche nächsten Schritte in der Entwicklung einer Technologie umgesetzt werden sollen. In dieser Interpre-

tation wird ein Aspekt erkennbar, der durchgehend in angedeuteter Weise mit dem Rollenkonzept in der CTA-Literatur verbunden ist: eine anklingende Kritik von fehlender Selbstreflexion und fehlender Aufmerksamkeit für die Bedarfe und Interessen der Menschen in einer Gesellschaft, die nicht über diese Informationen verfügt. Dies kann verschiedene bedenkliche Dynamiken hervorrufen, wie beispielsweise, dass bei der Gestaltung technischer Systeme eine zu einseitige Interessenslage vorherrscht, auf deren Grundlage Gestaltungsentscheidungen getroffen werden. Dies kann wiederum dazu führen, dass ein bestehendes Machtungleichgewicht noch stärker zur bereits schweren Seite ausschlägt. Um Technologien auf eine mehr ausgeglichene Weise zu gestalten, gilt es nach Ansicht in CTA, diesem Ungleichgewicht durch stärkeren Einbezug der Perspektiven außenstehender Akteure entgegenzuwirken. Verwendet wird das Rollenkonzept der Insider häufig auch als Synonym für die Impactors, die wiederum stellenweise als Synonym für Gestalter verwendet werden (siehe Kap. 5.4.4):

*„Technology developers are insiders and do not necessarily know very much about the ‘outside’. However, adoption and diffusion, is up to ‘outsiders’, who have other interests and expectations“* (Rip, 2008, S. 146).

Auch wenn die Bedeutungsparallelen der Insider zu den Impactors erkennbar sind, ist der Fokus bei den Insidern dennoch auf leicht andere Weise ausgerichtet. In CTA wird das Rollenkonzept in Verbindung mit dem Rollenkonzept der Outsider dazu verwendet, Gestaltungsprozesse als Teil gesellschaftlicher Vorgänge einzuordnen und mithilfe der durch sie zur Verfügung gestellten Unterscheidungsmöglichkeit zwischen einer Verortung außerhalb und innerhalb besonders die Grenzlinien zwischen diesen beiden gedachten Seiten nachzuzeichnen. Anders als bei den Impactors steht somit die gedachte oder tatsächlich existierende Verortung von Insidern im gesellschaftlichen Leben und besonders in Gestaltungsprozessen im Vordergrund. Durch das Ziehen solcher Grenzen können Insider innerhalb der gesellschaftlichen Kreise verortet werden, die die bedeutsamen, eine Technik prägende Entscheidungen treffen. Damit ist gelegentlich die Zuschreibung verbunden, dass sie diese Entscheidungen ohne eine CTA-Intervention primär vor dem Hintergrund fällen, um die eigenen Positionen weiter zu stärken und die eigenen Einschätzungen zu realisieren. Ausgehend von dieser Interpretation wird in CTA die Notwendigkeit abgeleitet, die Insider stärker auf die Lebensumstände und Interessen der Outsider aufmerksam zu machen und sie dazu zu bewegen, über die möglichen Konsequenzen ihrer Handlungen und Entscheidungen für Outsider zu reflektieren. Dabei werden Insidern in CTA nicht nur unterstellt, dass sie lediglich ihre eigenen Interessen und Ziele vor Augen haben. Es besteht auch die Möglichkeit, dass sie schlicht nicht von der Existenz der Outsider wissen und dass ihnen das notwendige Bewusstsein oder die notwendigen Mittel fehlen, sich mit ihnen auseinanderzusetzen. Aus dieser Möglichkeit wird in CTA ein weiteres Argument für die Notwendigkeit von CTA-Interventionen abgeleitet.

Insider zeichnen sich zudem dadurch aus, dass sie über Insider-Informationen verfügen (Wissensungleichgewicht), die ihnen einen strategischen Vorteil bieten. Aufgrund dieses Wissensvorsprungs wird ihnen in CTA punktuell eine besondere Verantwortung für

ihre nicht-informierten Mitmenschen zugeschrieben, um diese Informationen zur Schaffung von aus gesellschaftlicher Perspektive „better technologies“ einzusetzen. Ein weiterer Zweck den Insider und Outsider in CTA erfüllen ist auf einer etwas niedrigeren Abstraktionsebene und in einem anderen Betrachtungskontext, die an der Gestaltung technischer Systeme beteiligten wissenschaftlichen Disziplinen in Insider und Outsider aufzuteilen. So führen Rip und Robinson aus, dass nach ihrer Erfahrung viele der in Technikwissenschaften tätigen Menschen als Outsider des gesellschaftlichen Lebens betrachtet werden können, da sie in ihren Laboren in einer Art geschützten Raum leben und agieren und außerdem viel Aufwand darauf verwenden, diese (künstlichen) Grenzen zum gesellschaftlichen Geschehen aufrecht zu erhalten:

*„In a sense, scientists (even the technoscientists that abound in nanotechnology) are outsiders to society, because they live in protected places“* (Rip & Robinson, 2013, S. 50).

Dies stellt eine Kritik dar, denn in ihren weiteren Ausführungen machen Rip und Robinson deutlich, dass sie bei Menschen in geistes- und sozialwissenschaftlichen Tätigkeitsfeldern derlei Grenzen weniger erfahren haben. Dies führen sie auf den Umstand zurück, dass diese Menschen sich häufiger mit tatsächlichen sozialen Strukturen und Phänomenen im Feld auseinandersetzen (insertion). Sie schließen daraus, dass eine Öffnung der Technikwissenschaften hin zu einer höheren Reflexionsaktivität und -qualität dazu führen würde, aus gesellschaftlicher Perspektive „bessere“ TG zu fördern. Das Rollenkonzept der Outsider dient somit in CTA auch der Durchführung eines Perspektivwechsels **„nach innen“** in die an der Gestaltung von Technik beteiligten wissenschaftlichen Disziplinen. Durch das Aufzeigen der hier vorhandenen Grenzen können in CTA weitere Argumente für die in CTA vorgeschlagene methodology of insertion abgeleitet werden und dienen somit der Legitimierung von CTA.

## Enactors

Ein weiteres für CTA besonderes Rollenkonzept der informierten Einflussnehmer sind die Enactors mit dem ihnen als Gegenstück dienenden Rollenkonzept der comparative Selectors. Auch für Enactors gilt der zu Kapitelbeginn dargelegte übergeordnete Zweck, Gruppen und Individuen zu beschreiben, die auf häufig nicht näher definierte Weise Gestaltungsprozesse nach ihren Vorstellungen zu beeinflussen suchen. Die originäre Bedeutung der Bezeichnung enact (erlassen, verordnen, verfügen) deutet jedoch im Vergleich mit anderen informierten Einflussnehmern eine gewisse Präzisierung der Art ihrer Beeinflussung an, da es auf regulierende gesellschaftliche Instanzen schließen lässt.

Jedoch werden Enactors in CTA häufig als Technikgestalter interpretiert und als Synonym zu den Insidern verwendet (Rip & te Kulve, 2008, S. 66), weshalb einige der nachfolgend beschriebenen Charakteristika nicht von der Interpretation von Gestaltern in CTA getrennt werden können. Enactors werden beschrieben als

*„technology developers and promoters, who try to realize (enact) new technology, construct scenarios of progress and identify obstacles to be overcome“* (Rip & te Kulve, 2008, S. 52),

die demnach eine Funktion von Förderern von und (Be-)Werbern technischer Systeme erfüllen. Es kommt die Zuschreibung zum Ausdruck, dass sich Enactors aktiv für die Schaffung von Akzeptanz für Technik bei ihren Mitmenschen einsetzen. Diese Tätigkeit richtet sich an mögliche Technikskeptiker und Technikkritiker (siehe Kap. 5.4.8), die sie von den Vorteilen und Funktionen eines technischen Systems zu überzeugen suchen, in deren Gestaltung sie involviert sind oder von dessen Realisierung sie sich positive Effekte versprechen. Diese Funktion der Leistung von Überzeugungsarbeit bzw. der Schaffung von Akzeptanz klang bereits zuvor beim Rollenkonzept der Insider an, wird jedoch im Zusammenhang mit den Enactors noch deutlicher zum Ausdruck gebracht. Enactors sind so sehr überzeugt von dem (Zwischen-)Ergebnis eines Gestaltungsprozesses, dass seine positiven Effekte ihrer Ansicht nach auch für andere Gruppen und Individuen offensichtlich sein müssen und dass sich deshalb eine Betrachtung und Diskussion bestimmter technischer Funktionen oder Verfahren erübrigt. Aus Sicht der Enactors kann deshalb direkt zum nächsten Schritt übergegangen werden, der Suche nach konkreten Möglichkeiten zur Umsetzung und Weiterförderung eines technischen Systems.

Überzeugt von ihrer Position „Pro-Technik“ können sie jedoch blind für mögliche negative Effekte oder berechtigte Vorbehalte ihrer Mitmenschen sein. Diese Kritik wird in der CTA-Literatur verschiedentlich deutlich, weshalb das Rollenkonzept ebenfalls den Zweck erfüllt, Enactors einen Mangel an Reflexionsarbeit und -qualität zuzuschreiben. Stellenweise werden die Verhaltensweisen der Enactors auch als „ungestüm“ bewertet, da sie dazu neigen, die Positionen anderer Gruppen und Individuen zu ignorieren, zu delegitimieren oder gar direkt dazu übergehen, Argumente gegen ihre Position im Sinne einer Überstimmung (override) aus dem Weg schaffen. Diese Kritik steht im Zentrum der Bedeutung von enact wie sie in CTA interpretiert und gedeutet wird – die Neigung, sich auf inhaltlicher Ebene nicht mit begründeten Gegenargumenten auseinanderzusetzen.



Diese Tendenz wird in CTA als bedenklich eingeschätzt und dient daher auch als Argument für die in CTA vorgeschlagenen reflexionssteigernden Interventionsmaßnahmen. Welches Konfliktpotenzial in CTA in der Begeisterung und wenig hinterfragten Überzeugung für Technik identifiziert wird, bringen Rip und Robinson an einer Stelle mit folgenden Worten zum Ausdruck:

*„Enactors [...] see society as a constellation of possible barriers which have to be overcome. If questions are raised about the technology, such an enactor perspective will immediately see them as indications of potential barriers, even when the questions are mainly inquiry rather than criticism“ (Rip & Robinson, 2013, S. 43).*

Enactors werden durch Positionen irritiert, die die für sie auf der Hand liegenden Vorteile eines bestimmten technischen Systems oder einer Technologie nicht erkennen und in Frage stellen. Sie sind häufig sehr stark in einen Gestaltungsprozess investiert, weshalb sie direkt dazu übergehen, gegenläufige Argumente ihrer Position zu disqualifizieren:

*„This includes a tendency to disqualify opposition as irrational or misguided, or following their own agendas. [...] Enactors will get irritated, because for them, explaining the promise of their technological option should be enough to convince consumers/citizens“ (Rip & te Kolve, 2008, S. 52).*

Eine weitere Kritik, die in CTA aufgrund ihrer mangelnden Reflexionsarbeit mit Enactors verbunden wird, ist ihre geringe Bemühung um die kontinuierliche Auseinandersetzung mit und Suche nach alternativen Gestaltungsmöglichkeiten. Enactors setzen sich häufig aufgrund ihrer Befürwortung des Status Quo und den sich abzeichnenden weiteren Entwicklungen eines Gestaltungsprozesses kaum noch mit möglichen Alternativen in der Gestaltung und Nutzung von Technik auseinander. Dadurch neigen sie dazu, die sich andeutenden weiteren Entwicklungen als gegeben anzunehmen und in bereits beschlossene Entwicklungspfade zu übersetzen:

*„A structural problem is that enactors tend to project a linear future, defined by their intentions, and use the projection as a roadmap - only to be corrected by circumstances“ (Rip, 2018, S. 104).*

Da sie die weiteren Verläufe eines Gestaltungsprozesses identifiziert zu haben glauben, verlieren sie ihre gedankliche und operative Flexibilität und neigen zu einer gewissen Form von Engstirnigkeit, die sich später aus Sicht in CTA schwer rächen kann. Gerade in frühen Phasen von Gestaltungsprozessen wird in CTA ein hoher Wert in der Schaffung und Aufrechterhaltung eines gedanklichen Kontingents für alternative Entwicklungen oder unerwartet auftretender Ereignisse identifiziert. Entsprechend dient diese Eigenschaft des Rollenkonzepts der Enactors auch dazu, die CTA-Position für eine längerfristig angelegte Reflexionsarbeit in Gestaltungsprozessen zu stützen. Ob Enactors über formell anerkannte, tatsächliche Befugnisse eines steuernden Typs verfügen, bleibt in der CTA-Literatur weitgehend unklar. Enactors besitzen zudem nach Interpretation in CTA eine konzentrische Perspektive auf ihre Umwelt, bei der ein Gestaltungsprozess den Mittelpunkt dar-

stellt und um den sich verschiedene weitere Schichten (layers) des gesellschaftlichen Lebens befinden, an deren Eigenschaften und Bedingungen ein technisches System im Laufe eines Gestaltungsprozesses angepasst werden muss. Die Entfernungen der verschiedenen Schichten zum Mittelpunkt stellen inhaltlich und räumlich die Entfernung zu typischen Inhalten und Fragestellungen eines Gestaltungsprozesses dar, wodurch nach Einschätzung in CTA beispielsweise die Ziele eines nahe am Mittelpunkt gelegenen und agierenden Unternehmens häufiger und intensiver in einem Gestaltungsprozess thematisiert werden, als die Fragen und Auswirkungen, die in vergleichsweise weiter entfernten und übergeordneten gesellschaftlichen Strukturen zum Vorschein treten. Ausgehend von dieser Zuschreibung einer konzentrischen Perspektive wird in CTA ferner davon ausgegangen, dass sich Enactors darum bemühen, die durch die verschiedenen Schichten notwendigen Anpassungen eines technischen Systems sukzessive zu realisieren, jedoch orientieren sie sich dabei von innen nach außen und priorisieren nähergelegene Aspekte höher wie weiter entfernte. Im Zusammenhang dieser als Teil eines Gestaltungsprozesses sukzessiven Bewegung nach außen wird Enactors in CTA eine sukzessive Vernachlässigung und geringere Priorisierung unterstellt, weshalb die für CTA bedeutsamen gesellschaftlichen Fragestellungen und Bedenken deutlich seltener, deutlich später und deutlich weniger intensiv in Gestaltungsprozessen thematisiert werden, wie jene Aspekte, die sich auf die notwendigen Anpassungen näher gelegener Schichten beziehen. Somit dient das Rollenkonzept in CTA auch dazu, die Priorisierung gesellschaftlich relevanter Fragestellungen in Gestaltungsprozessen zu beschreiben und ihre geringe Zentrierung zu kritisieren.

## Comparative Selectors

Die Charakterisierung der comparative Selectors in CTA nimmt stark Bezug zum Rollenkonzept der Enactors, zu dem an entscheidenden Stellen Ausprägungsunterschiede deutlich gemacht werden. In CTA wird Enactors zugeschrieben, dass sie ihre Mitmenschen häufig als Technikempfänger sehen, die darauf warten, ein technisches System (endlich) zum empfangen. Die Enactors selbst leiten daraus für sich selbst die Funktion als eine Art Wegbereiter oder Orientierungshelfer ab, um den Empfangsprozess bestmöglich zu unterstützen. Die CTA-Literatur weist nun darauf hin, dass sich die soziale Realität häufig entscheidend von dieser Einschätzung unterscheidet, da es in Gestaltungskontexten häufig gesellschaftliche Instanzen gäbe, die durchaus in der Lage seien, Alternativen zu den Gestaltungsmöglichkeiten eines betreffenden technischen Systems oder einer Technologie in Erwägung zu ziehen. Diese als comparative Selectors bezeichneten Instanzen vergleichen zwischen verschiedenen technischen Möglichkeiten und Anwendungen, wägen ihre Vor- und Nachteile ab und entziehen sich damit dem im Rollenkonzept der Enactors in Anspruch genommenen Vormachtstellung (aufgrund von Wissensvorteilen) und der damit verbundenen Tendenz einer Bevormundung. Comparative Selectors gehen bewusst gedanklich einen oder mehrere Schritte zurück und setzen sich mit möglichen oder notwendigen Alternativen auseinander, und teilen deshalb auch selten die hohe und ungeteilte Überzeugung für ein bestimmtes Gestaltungs- und Nutzungsverfahren. Damit ist das zentrale Merkmal der comparative Selectors dargelegt, denn sie vergleichen und bewer-

ten verschiedene Gestaltungsalternativen anhand Kriterien, die für einen Gestaltungskontext relevant sind:

*„[...] and “comparative selectors” who can consider different options to select from and do formal or informal versions of cost-risk-benefit assessment” (Rip & Robinson, 2013, S. 43).*

Dazu analysieren sie relevante Kennzahlen und Indikatoren um beispielsweise Statistiken zu erstellen und legen dadurch den Fokus häufig auf relevante Fragen, die in den von Enactors vorgeschlagenen und vertretenen Gestaltungs- und Weiterentwicklungspfaden wenig Beachtung finden oder als vernachlässigbar eingestuft werden. Sie gründen ihre Entscheidungen auf das Wissen, das diese Aktivitäten zum Vorschein bringen, und mit dessen Hilfe sie dann wiederum Einfluss auf Gestaltungsprozesse ausüben.

Bei comparative Selectors kann es sich auf einer konkreteren Ebene um gesellschaftliche Gruppen handeln, deren gesellschaftlicher Mehrwert darin besteht, dass sie genau solche Vergleiche durchführen. So nennt die CTA-Literatur die US Food and Drug Administration als ein Beispiel für einen comparative Selector, da sie zur Aufgabe hat, Arznei- und Lebensmittel auf Wirksamkeit und Gefahrenpotenzial zu untersuchen. Zwischen Enactors und comparative Selectors können entsprechend sehr formelle Beziehungen bestehen, da Enactors teilweise gesetzlich dazu verpflichtet sind, mit comparative Selectors zusammenzuarbeiten und sich ggf. deren Urteil oder ihren Vorgaben zu beugen. Somit ist es durchaus möglich, dass comparative Selectors in CTA aufgrund ihrer gesellschaftlichen Verortung weisungsbefugt gegenüber Enactors sind, und diese zur Realisierung bestimmter Entwicklungspfade oder zur Korrektur bereits eingeschlagener Entwicklungspfade verpflichten können. Auf diese Beziehungsebene verweist auch die Promoter-Funktion der Enactors, denn sie werben für ihre Produkte und Überzeugungen auch in solch formellen Interaktionsbeziehungen mit comparative Selectors (siehe Kap. 5.4.7). Jedoch weist die CTA-Literatur auch darauf hin, dass zwischen beiden Seiten auch mehr informelle Interaktionsbeziehungen bestehen können, wie beispielsweise zwischen Managementabteilungen von Unternehmen und gesellschaftlichen Vertretern wie Vereinen oder Umweltgruppierungen. Dementsprechend besteht in CTA durchaus die Möglichkeit, dass auch Bürgerinnen und Bürger als comparative Selectors betrachtet werden:

*„There are also citizens – consumers, etc., as amateur comparative selectors – which can range more freely because they are not tied to certain methods and to accountability” (Rip & te Kulve, 2008, S. 52).*

Demnach zeichnen sich „laienhafte“ comparative Selectors durch die besondere Eigenschaft aus, dass sie im Vergleich zu professionellen comparative Selectors deutlich häufiger und direkter Gegenargumente zu einer bestimmten Entwicklung zum Ausdruck bringen, die sich aus ihrer derzeitigen, alltäglichen Lebenssituation speist.

Es ist nicht von der Hand zu weisen, dass im Rollenkonzept der comparative Selectors Institutionen repräsentiert werden, die TA-Arbeit verrichten. Im Vergleich und in der Bewertung technologischer Alternativen besteht ein zentrales traditionelles Aufgaben-

feld von TA. Somit dient das Rollenkonzept in CTA auch dazu, viele in TA übliche Argumentationslinien zur Technologiebewertung anzubringen und stützt dadurch zusätzlich das Plädoyer für das Vorhalten einer gewissen gedanklichen Offenheit und Flexibilität, um dem voreiligen Ausschluss von Gestaltungsalternativen entgegenzuwirken. Diese Vergleichsarbeit wird auch durch die in CTA vorgeschlagenen Methoden angestrebt, besonders vor dem Hintergrund der Identifikation der von technologischen Entwicklungen häufig negativ Betroffenen. Zusätzlich zur Gegenpolfunktion zu den Enactors bietet das Rollenkonzept der comparative Selectors in CTA die Möglichkeit, TA-Arbeit in Gestaltungsprozessen zu verorten.

### 5.4.3 Die Unbeteiligten und Betroffenen

Personen, die der Rollenkonzeptgruppe der Unbeteiligten und Betroffenen zugeschrieben werden, werden in CTA und VSD am deutlichsten dadurch charakterisiert, dass ihnen die Möglichkeit zur Partizipation und Mitbestimmung in Gestaltungsprozessen nur in sehr geringem Umfang möglich ist oder gänzlich verwehrt ist. Sie sehen sich häufig mit den Ergebnissen der in Gestaltungsprozessen gefällten Entscheidungen und den daraus resultierenden Systemen konfrontiert und müssen ihre Adaptionfähigkeit unter Beweis stellen, um die gestalteten technischen Produkte in ihre Lebens- und Arbeitssituationen zu integrieren, um mit ihrer Hilfe wiederum die von ihnen gewünschten oder geforderten Tätigkeiten durchführen zu können. Somit steht bei den hier thematisierten Rollenkonzepten, wie auch bei den informierten Einflussnehmern die Machtfrage im Vordergrund. Der Aspekt der Beteiligung bezieht sich dabei besonders auf das aktive Gestaltungsvorgehen, während das Betroffensein einen Gestaltungsprozess zeitlich vom Ende her betrachtet. Somit bilden die Unbeteiligten und Betroffenen (mit Ausnahme der Vorbereitungsphase) den gesamten zeitlichen Verlauf eines Gestaltungsprozesses ab, in dem sie zu unterschiedlichen Zeitpunkten und auch wiederholt relevant sein können. Auch dienen sie der Legitimation, da in beiden Ansätzen die Vorzüge und die Notwendigkeit partizipativer Gestaltungsverfahren hervorgehoben werden. In CTA und VSD gilt es, die unbeteiligten Gruppen und Individuen, die für einen Gestaltungsprozess Relevanz besitzen zu identifizieren und zu beteiligen, um ihre Sichtweisen und Anforderungen in Gestaltungsprozessen berücksichtigen zu können und um dadurch negativen Folgewirkungen vorzubeugen, die wiederum viele Betroffene zum Ergebnis haben können. Der sprachliche Konjunktiv ist dabei bezeichnend für die Betroffenen, denn sie besitzen im Gestaltungszusammenhang häufig einen antizipierten Charakter, da sich seine Folgen häufig nur mithilfe der Formulierung von Annahmen beschreiben lassen.

Besonders in CTA werden die Unbeteiligten und Betroffenen als die Verlierer in der Gewinner-Verlierer-Analogie von Gestaltungsprozessen ausgelegt. In den Argumentationsstrukturen in VSD und CTA sind die Rollenkonzepte der Unbeteiligten und Betroffenen von zentraler Bedeutung, da sie als Benchmark für eine sozial- und werteorientierte Gestaltung fungieren. Weist ein Gestaltungsprozess nur geringe Partizipationsmöglichkeiten oder hohe Risiken für eine große Anzahl negativ Betroffener auf, wird in beiden Ansätzen die Position vertreten, dass sozialen und ethischen Aspekten nicht die ihnen angemessene Aufmerksamkeit gewidmet wird. Die Abwesenheit entsprechender

Vorkehrungen dient in beiden Ansätzen als Ansatzpunkt, um die in ihnen vorgeschlagenen Methoden in Gestaltungsprozesse einzubringen. Die Rollenkonzepte der Unbeteiligten und Betroffenen stellen auch ein wichtiges Bindeglied zur MZG dar, da auch hier partizipative Methoden präferiert werden und ebenfalls eine möglichst vollständige Ermittlung menschlicher Bedürfnisse und Werte im Vordergrund steht. In VSD werden die Unbeteiligten und Betroffenen hauptsächlich mit dem Rollenkonzept der indirect Stakeholders abgehandelt, während sich mit ihnen in CTA überwiegend mit dem Rollenkonzept der Impactees auseinandergesetzt wird.

## Indirect Stakeholders

Indirect Stakeholders zeichnen sich in VSD dadurch aus, dass sie nicht dazu in der Lage sind, mit einem System zu interagieren. Durch die Betrachtung der indirect Stakeholders wird in VSD ein Gestaltungsprozess bewusst vom zeitlichen Ende aus betrachtet, um indirekte Stakeholder identifizieren zu können. Dies soll wiederum ermöglichen, die Werte der indirekten Stakeholder zu identifizieren um sie durch einen erneuten Perspektivwechsel in der Gestaltungsgegenwart berücksichtigen und adressieren zu können. In VSD werden indirekte Stakeholder bewusst in den Mittelpunkt der Aufmerksamkeit gerückt, da mit ihnen eine in VSD vertretene, grundsätzliche Kritik an vielen Gestaltungsprozessen und Gestaltungsframeworks verbunden ist. Demnach ignorieren viele Gestaltungsverantwortliche die mittelbar betroffenen indirect Stakeholders, da sie scheinbar vernachlässigbare Relevanz für den Gestaltungsfortschritt und Erfolg besitzen. Zu dieser Einstellung wird in VSD mit den indirect Stakeholdern bewusst ein Gegenargument präsentiert. Auch wird anerkannt, dass es sich häufig als schwierig darstellt, indirekte Stakeholder aufgrund ihrer mittelbaren Beziehung zum Gestaltungsprozess zu identifizieren. Da indirekte Stakeholder nicht direkt mit einem System interagieren, erfordert ihre Identifikation eine ausführliche Auseinandersetzung mit dem Gestaltungskontext in dem bzw. für den ein System gestaltet wird. Dies erfordert häufig einer Erweiterung der gedanklichen Grenzen, in denen sich Gestaltungsverantwortliche typischerweise bewegen (B. Friedman & Hendry, 2019, S. 42), worin die Schwierigkeit der Identifikation begründet liegt. In VSD wird jedoch nicht davor zurückgescheut, diesen zusätzlichen Aufwand zu betreiben, da die mit ihm assoziierten Resultate den Mehraufwand ausgleichen und rechtfertigen. Durch das Befolgen der tripartite methodology soll es Gestaltungsverantwortlichen ermöglicht werden, die dafür notwendige gedankliche Grenzerweiterung auf möglichst niederschwellige Art zu vollziehen. Das Rollenkonzept der indirect Stakeholders dient in VSD somit einerseits dazu, die gedanklichen blinden Flecke in Gestaltungsprozessen zu beleuchten und besitzt zusätzlich die Funktion, in einem Gestaltungsprozess ein möglichst ganzheitliches Verständnis über den Gestaltungskontext zu erarbeiten. Im Vergleich mit den Interpretationen der Impactees in CTA besteht in VSD die Möglichkeit, dass indirekte und unbeabsichtigte Folgen von Gestaltungsprozessen nicht nur negative, sondern durchaus auch positive Effekte bewirken können. Somit besteht in VSD eine leicht positivere Einstellung gegenüber den Unbeteiligten und Betroffenen.

Neben den mehr übergreifenden Attributen werden indirect Stakeholders auch konkreter beschrieben. Ein typisches Beispiel für eine konkrete Ausprägung ist ein „Bystander“ (B. Friedman & Hendry, 2019, S. 41), der beispielsweise im Kontext eines Telefongesprächs im öffentlichen Raum indirekt durch die Lautstärke des Telefonats betroffen ist und in der er z.B. in einer U-Bahn nur geringe Möglichkeiten hat, sich den Auswirkungen dieser Interaktionssituation zu entziehen. Als ein weiteres Beispiel werden in der Literatur Zuschauer oder Passanten im Straßenverkehr angeführt, die von Kameras erfasst werden, die Polizeibeamte an ihrer Arbeitskleidung tragen. Auch in diesem Kontext besitzen die Passanten nur wenig direkte Einflussmöglichkeiten auf die Nutzung einer Technologie, jedoch besitzen sie Werte, die nach Auffassung in VSD in der Gestaltung solcher Systeme berücksichtigt werden sollten. Ein letztes Beispiel für indirekte Stakeholder sind ungeborene Kinder. Sie können keinen Einfluss darauf nehmen, wie ihre Eltern beispielsweise Social Media-Anwendungen nutzen, können aber später in ihrem Leben von diesen Verhaltensweisen sowohl profitieren als auch Nachteile erlangen. Mit solchen Beispielen wird in der VSD-Literatur verdeutlicht, weshalb es so schwierig sein kann, angemessene Lösungen und Kompromisse im Umgang mit indirekten Stakeholdern zu finden. Häufig besteht nur die Wahl zwischen zwei unliebsamen Alternativen. Ein weiterer Mehrwert, den solche Beispiele in VSD bieten, ist die Verdeutlichung der interactional stance. In vielen der beispielhaften Interaktionssituationen ist durchaus vorstellbar, dass im Beispiel des Telefonats im öffentlichen Raum eine Person in einer Situation direkter Stakeholder ist (Telefonat führen) während in einer anderen Situation dieselbe Person den indirekten Stakeholder darstellt (Passant in U-Bahn). Derartige zusätzliche Perspektivwechsel ermöglichen es in VSD, weitere relevante Gestaltungsaspekte freizulegen und fördern zudem die Empathie von Gestaltungsverantwortlichen, da sie mit hoher Wahrscheinlichkeit solche alltäglichen Situationen selbst einmal erlebt haben und im Übertrag auf ihren Gestaltungskontext nachvollziehen können, welche Relevanz die Berücksichtigung indirekter Stakeholder besitzen kann.

Da in VSD konkrete Interaktionskontexte zur Identifikation indirekter Stakeholder verwendet werden, besteht eine weitere Funktion des Rollenkonzepts darin, die Zielsetzungen eines Gestaltungsprozesses auf nicht-intendierte Auswirkungen zu überprüfen. In vielen Gestaltungsprozessen sehen sich Verantwortliche mit dem Dilemma konfrontiert, neben den intendierten Zwecken eine häufig größere Zahl nicht-intendierter Zwecke zu bewirken. Das Rollenkonzept der indirect Stakeholders ermöglicht es, sich mit letzteren intensiv auseinanderzusetzen. In VSD werden indirect Stakeholder deshalb stellenweise auch als eine „non-targeted role“ (Nathan et al., 2008) beschrieben. Sie ermöglichen es, untypische Anwendungsfälle zu beschreiben, die deshalb untypisch sind, da sie von den Gestaltungsverantwortlichen nicht mit Intention geschaffen werden. Konkrete Ausprägungen solcher indirect Stakeholder werden in der Literatur als Angestellte beschrieben, die Wege finden, um vorgesehene Arbeitsschritte eines Systems durch findige Manipulationen zu umgehen. Aus diesem Grund sind indirekte Stakeholder potenziell auch für Sicherheitsfragen in Gestaltungsprozessen relevant, da solche Workarounds in einem System auch für mögliche Hackingangriffe verwendet werden können. Als eine weitere mögliche Ausprägung von indirect Stakeholdern führen Friedman und Hendry (2019, S. 42) den sogenannten „human data point“ an, womit sie sämtliche Situationen des gesellschaftlichen Lebens bezeichnen, in denen in IT-Systemen Daten über Menschen erhoben

und erfasst werden, ohne den Menschen die Möglichkeit zu geben, auf diese Daten zuzugreifen. Patientensysteme, Überwachungssysteme, Automobilsensoren oder auch Systeme, die den Wasserverbrauch eines Haushalts erfassen, zählen sie zu solchen Systemen.

Das Rollenkonzept der indirect Stakeholders dient in VSD in Summe dazu, den Betrachtungshorizont der bereits identifizierten Stakeholder auf möglicherweise weitere relevante Stakeholdergruppen auszuweiten. Auch können mit ihm weitere für Gestaltungsprozesse relevante Werte identifiziert werden und Erkenntnisse über bestehende Wertkonflikte unter Menschen erarbeitet werden. Das Rollenkonzept dient dazu, das Bewusstsein für die Eigenschaften und Besonderheiten eines Gestaltungskontexts zu schärfen (Sensitivität) und um bei Gestaltungsverantwortlichen ein Bewusstsein über die Tragweite ihrer Entscheidungen zu schaffen, über das sie möglicherweise nicht verfügen. In VSD wird dabei nicht der Anspruch erhoben, immer die zutreffenden und angemessenen Grenzen für einen Gestaltungsprozess definieren zu können, es wird jedoch auf die Notwendigkeit der Erweiterung verwiesen, wenn die Zielsetzung lautet, Systeme wertsensitiv zu gestalten: „Identifying and documenting other types of indirect stakeholder roles remains an open line of inquiry“ (B. Friedman & Hendry, 2019, S. 42).

## Impactees

Die Impactees stellen das CTA-Pendant zu den indirekten Stakeholdern in VSD dar, weshalb ich an dieser Stelle lediglich auf die Bedeutungsabweichungen zwischen beiden Rollenkonzepten eingehe. Im Gegensatz zu ihrem Gegenstück-Rollenkonzept, den Impactors, verfügen Impactees über deutlich weniger Wissen über technische und strukturelle Sachverhalte und sind ebenso wie die indirekten Stakeholder typischerweise deutlich seltener aktiv in Entscheidungsprozesse eingebunden. Im Unterschied zu den indirekten Stakeholdern tritt bei den Impactees die negative Konnotation deutlicher zum Vorschein, da sie in CTA deutlich klarer als die Verlierer von Gestaltungsprozessen betrachtet werden: „Who represents actual or potential users, victims or impactees in CTA activities, and how?“ (Schot & Rip, 1997, S. 265). In CTA steht der Aspekt nicht-intendierter Effekte von Technik, die Impactees erfahren, im Vordergrund. Es wird danach gestrebt, mithilfe des Rollenkonzepts nicht-intendierte Effekte zu identifizieren, zu benennen und ihre Auswirkungen Gestaltungsverantwortlichen vor Augen zu führen, um sie letztlich abzuschwächen oder gänzlich abwenden zu können. Das Rollenkonzept der Impactees dient dabei dazu, die Empfängerseite (receiving end) des wahrgenommenen oder tatsächlichen Machtungleichgewichts in Gestaltungsprozessen darzulegen und die Notwendigkeit für Veränderungen im Gestaltungsprozess zu verdeutlichen. Impactees soll deshalb eine deutlich stärkere Auseinandersetzung und Berücksichtigung gewidmet werden, wie es in Gestaltungsprozessen ohne CTA-Interventionen der Fall wäre. Impactees werden dabei in der CTA-Literatur selten mit weiteren Attributen versehen, eine der deutlichsten Charakterisierungen findet sich bei Parandian, der Impactees beschreibt als „professional and non-professional users, and broader groups that experience unintended secondary effects“ (Parandian, 2012, S. 34). Die Impactees stellen diejenige Gruppe von Akteuren dar, die in CTA dazu verwendet werden, um die als kritisch

eingestuften Verhältnisse des two-track approachs zu verdeutlichen, der in CTA durch die vorgeschlagenen Interventionen modernisiert werden soll. Somit repräsentieren die Impactees die Hauptzielgruppe vieler TA-Arbeiten – sie sollen nach Auffassung in TA in den Aufmerksamkeitsmittelpunkt von Gestaltungsprozessen gerückt werden und mit Verweis auf demokratische Prinzipien ein Mitspracherecht erhalten. Sämtliche Aktivitäten in CTA sind darauf ausgerichtet, Impactees zu identifizieren, ihre Relevanz zu verdeutlichen und ihre Positionen in Gestaltungsprozessen zu stärken, um entweder mit Verweis auf die Vergangenheit oder antizipierend in die Zukunft unerwünschte Nebenfolgen abzuwenden. „Bessere“ Technologien zeichnen sich nach Ansicht in CTA dadurch aus, dass sie eine geringe Anzahl an negativ betroffenen Impactees bewirken.

#### 5.4.4 Die Gestalter

Die Rollenkonzeptgruppe der Gestalter vereint sämtliche Rollenverständnisse, in denen Menschen als über technische oder kreativ-konzeptionelle Expertise verfügend konzipiert werden, die sie zum Zweck der Gestaltung technischer Systeme in Gestaltungsprozesse einbringen. Typische Vertreter die sich in dieser Gruppe wiederfinden sind die Rollenkonzepte der (Technology) Developers und der Engineers, sowie das besonders in VSD präsente Rollenkonzept der Designer. Die zentrale Funktion dieser Rollenkonzepte ist es, zum Ausdruck zu bringen, dass das Wissen, über das sie verfügen, zur Gestaltung technischer Systeme eingebracht und angewendet wird. Dadurch wird es in beiden Ansätzen möglich, auf einer etwas tieferen Abstraktionsebene sich konkreter mit tatsächlichen Gestaltungssettings auseinanderzusetzen, um sich den Spezifika eines Gestaltungskontexts anzunähern und eine grundsätzliche Aufteilung der involvierten Gruppen und Individuen vorzunehmen. Typischerweise werden hierfür die Rollenkonzepte der Gestalter den Rollenkonzepten der Nutzer (siehe Kap. 5.4.6) gegenübergestellt, um die zwischen ihnen bestehenden Abhängigkeiten, Wechselwirkungen und auch Unterschiede zu beschreiben und zu analysieren: „the dichotomies (innovation vs. responsible, technology developers vs. users) remain visible“ (Rip & Robinson, 2013, S. 41). Diese gedankliche Aufteilung zwischen Gestaltung und Nutzung ist in der Literatur zu beiden Ansätzen durchgehend erkennbar und kommt besonders in VSD zum Tragen, in CTA wird auf dieser Abstraktionsebene stellenweise noch der zusätzliche dritte Akteur der Regulierung betrachtet. Im Vergleich mit den bisher dargelegten Rollenkonzepten wird deutlich, dass die Rollenkonzepte der Gestalter sich deutlich stärker an konkreten Berufsbildern orientieren, wodurch die etwas geringere Abstraktionshöhe unterstrichen wird. Dies wird auch deutlich beim Rollenkonzept der Designer, das sich häufig auf die Charakteristika von realen Berufsbildern im Produktdesign bezieht. Die CTA- und VSD-Literatur verwendet die Rollenkonzepte der Gestalter häufig synonym, eine bewusst durchgeführte Unterscheidung zwischen Ingenieuren und Entwicklern findet selten statt. Ebenso selten werden in den Ansätzen zusätzliche Präzisierungen hin zu existierenden Fachrichtungen innerhalb gestaltender Disziplinen (z.B. Maschinenbau, Lebensmittelchemie, Architektur o.ä.) durchgeführt. Dies würde der in Anspruch genommenen Übertragbarkeit zwischen Gestaltungskontexten entgegenwirken. Dennoch sind diesbezüglich zwischen VSD und CTA Nuancen erkennbar, denn in VSD werden technische Gestalter häufig im Kontext der



Gestaltung von IT-Systemen betrachtet, während in CTA hier durchaus breitere Betrachtungsspektren Anwendung finden. Punktuell wird in der Literatur beider Ansätze auch die Bezeichnung der Researcher mit den bereits genannten Rollenkonzepten gleichgesetzt, ohne diese jedoch näher zu erläutern. Dies erweckt den Eindruck, dass in beiden Ansätzen durchaus auch Forscher als Gestalter betrachtet werden. Dies zeigt eine gewisse Selbstreflexion der Autoren und deutet zudem an, eine Relevanz für Forschungstätigkeiten innerhalb und außerhalb staatlich geförderter Forschung zu besitzen.

## Developers und Engineers

Während die technisch gestaltenden Rollenkonzepte stellenweise mit dem Rollenkonzept der Engineers bezeichnet werden, ist in CTA und VSD das Rollenkonzept der Developers deutlich stärker vertreten. Es bezeichnet in beiden Ansätzen Gruppen und Individuen, die über Fachwissen in technischen Domänen verfügen und dieses Wissen zur Gestaltung technischer Systeme anwenden. In VSD werden hierbei häufig Parallelen zur Bedeutung von Softwareentwicklern erkennbar (z.B. Lüthi et al., 2021). So wird das Rollenkonzept in VSD häufig mit entsprechend typischen Charakteristika assoziiert, wie beispielsweise Programmiertätigkeiten mithilfe verschiedener Programmiersprachen. Auch weitere für derartige Berufsbilder typische Anforderungen und Fähigkeiten werden in der Literatur referenziert, wenn beispielsweise die Integration eines Systems in eine bestehende Systemlandschaft oder die Wartung eines bestehenden Systems erfolgen soll. Daran wird eine weitere Funktion der Rollenkonzepte erkennbar, denn es ermöglicht in beiden Ansätzen, verschiedene Aufgaben- und Verantwortungsbereiche zu unterscheiden und zu analysieren. In VSD ist dies von Interesse, da es dadurch möglich wird, die Wertefragen betreffenden Aspekte und Tätigkeiten von Entwicklern und Ingenieuren zu identifizieren und sich intensiv mit ihnen auseinanderzusetzen. So heben beispielsweise Lüthi et al. (Lüthi et al., 2021, S. 195) hervor, dass häufig ein grundsätzlicher Unterschied darin besteht, wie Developer die Werte von Nutzern interpretieren, und wie Nutzer selbst ihre Werte interpretieren. Aus diesen unterschiedlichen Interpretationsweisen entstehen häufig die in VSD zentralen Wertekonflikte (value tensions), die es gilt transparent werden zu lassen und zu adressieren. In VSD werden entsprechend viele Ressourcen vorgesehen, um sich mit sämtlichen Aspekten auseinanderzusetzen, die Aufschluss über die Werte von Developern, ihren Vorgehensweisen zur Werteinterpretation und ihrem Umgang mit Wertveränderungen in einem Gestaltungsprozess beschäftigen. So beschreibt beispielsweise Korenhof (2023), dass Softwareentwickler bei ihrer Entscheidungsfindung in Gestaltungsprozessen verschiedene Faktoren berücksichtigen, die sich zusammensetzen aus ihrem Fachwissen, ihrer Praxiserfahrung und ihrem kulturellen Hintergrund. Dabei nehmen sie unweigerlich subjektive Interpretationen ihrer Aufgabenstellung und der Zielsetzung einer Software vor, die sie bei der Durchführung von Programmiertätigkeiten in Softwarecode transformieren. Solche Darlegungen geben Aufschluss darüber, welche Komplexität in Gestaltungsprozessen auf der Ebene von Wertefragen präsent sein kann und wie schnell sich entsprechend auch Betrachtungen in VSD mit einem hohen Grad an Komplexität konfrontiert sehen, die es angemessen zu adressieren gilt. Die Rollenkonzepte der Gestalter ermöglichen es, eine tiefergreifende und konkrete Auseinandersetzung

mit den komplexen Zuständen und Abhängigkeiten gestalterischen Handelns durchzuführen, die neuralgischen, Wertefragen betreffende Aspekte und Tätigkeiten eines Gestaltungsprozesses zu identifizieren und zu analysieren und besonders die wertenden Elemente gestalterischen Handelns hervorzuheben:

*„assumptions and biases are incorporated in the software – whether it be intentional or accidental [...]. Coded procedures are therefore inevitably value laden“*  
(Korenhof, 2023, S. 505).

Solch tiefreichende Betrachtungen der Handlungen und Anforderungen von Ingenieuren und Entwicklern dienen in beiden Ansätzen auch dazu, die zentrale Annahme zu unterstützen, die Gestaltung von Technik nicht als wertneutrale oder wertfreie Tätigkeit zu betrachten. Das Aufzeigen der wertenden Elemente gestalterischen Handelns dient dabei häufig als zentrales Argument, jedoch werden Developer in beiden Ansätzen auch unter dem Aspekt ihrer Machtposition betrachtet. Besonders in CTA wird hierfür auf Gedanken zurückgegriffen, die dort auch im Zusammenhang mit den Impactors (siehe Kap. 5.4.2) angeführt werden. Jedoch wird auch in der VSD-Literatur darauf hingewiesen, dass Developer, wenn sie beispielsweise über eine Systemarchitektur entscheiden, maßgeblichen Einfluss auf die Möglichkeiten nehmen, die Nutzern Zu- und Umgang mit einem System erlauben. Somit steht das Rollenkonzept der Developer und Engineers in der VSD- und CTA-Literatur durchgehend mit dem Aspekt der Verantwortung in Verbindung, die Developer und Engineers nach Ansicht in CTA und VSD gegenüber ihren Mitmenschen besitzen. Viele Arbeiten besonders zu VSD setzen sich mit den verschiedenen Facetten auseinander, die diese Verantwortung ausmachen können. Nicht selten stehen die Gefahren im Vordergrund, die mit einem lässigen oder unbedachten Umgang mit dieser Verantwortung verbunden sein können. In VSD wird dies in einer Arbeit am Beispiel von Nudging dargestellt, dass das Phänomen beschreibt, wenn Menschen durch bewusst oder unbewusst gestaltete Systemeigenschaften auf subtile Weise zu Verhaltensweisen und Verhaltensänderungen motiviert werden (Asikis et al., 2021). In beiden Ansätzen wird versucht, durch die Verwendung der Rollenkonzepte der Developer und Engineers sowie mithilfe der damit zusammenhängenden vorgeschlagenen Methoden, Entwickler und Ingenieure auf ihre Verantwortung aufmerksam zu machen und ihnen Wege aufzuzeigen, wie sie dieser Verantwortung gerecht werden können. Besonders in VSD dienen die Rollenkonzept auch dazu, Wissen und Verständnis für die Arbeitsbedingungen und Herausforderungen der Developer und Ingenieure zu demonstrieren (progress, not perfection).

Während in VSD das Rollenkonzept der Developer und Engineers in starker Anlehnung zum IT-Kontext verwendet wird, wird in CTA mehrheitlich von den übergeordneten „Technology Developers“ gesprochen (Rip & Robinson, 2013; Rip & te Kulve, 2008; Schot & Rip, 1997). Hier wird sich deutlich seltener auf bestimmte technische Domänen wie beispielsweise die Elektrotechnik oder die Informatik bezogen, sondern es herrscht eine deutlich allgemeinere Vorstellung zu den Fähigkeiten, Handlungen und Verantwortungsbereichen der Developer und Engineers vor. Auch in CTA wird mit der grundlegenden Unterscheidung zwischen Gestaltung und Nutzung gearbeitet, allerdings erfolgt dies hier

vor dem Hintergrund, die Notwendigkeit zu verdeutlichen, Gestaltungsprozesse differenzierter zu betrachten. In CTA wird die Position vertreten, Innovationsprozesse als sich wechselseitig beeinflussende, ko-evolutionäre Aushandlungsprozesse verschiedener sozialer Akteure darzustellen. Damit in diesen komplexen sozialen Strukturen eine Aussicht auf Erfolg im Sinne der Gestaltung sozial akzeptabler Technologien bestehen kann, ist nach Ansicht in CTA in Gestaltungsprozessen ein „mutual learning“ unter sämtlichen beteiligten Akteuren notwendig. Hierbei wird besonders an die Technology Developers appelliert, die sich mehr den Gedanken, Bedarfen und Vorbehalten sozialer Akteure zuwenden sollen (CTA-Workshops). Diese Notwendigkeit wird damit verargumentiert, dass moderne Technologien nicht nur auf technischer Ebene, sondern auch auf sozialer Ebene deutlich höhere Komplexitätsgrade erreichen. Als Beispiele hierfür dienen die Nanotechnologie, die synthetische Biologie oder das Geo-Engineering. Dies ermöglicht in CTA auch aufzuzeigen, dass aufgrund dieser höheren Komplexität der häufig übliche Usus der Unterteilung in Gestaltung und Nutzung nicht mehr angebracht ist, da er nicht in der Lage ist, die vorhandene Komplexität abzubilden: „The simple contrast between technology developers and users is inapplicable then“ (Rip & Robinson, 2013, S. 42). Darin besteht eine Parallele zur Verwendung des Rollenkonzepts in VSD, denn auch dort wird durch die Betrachtung der moralischen Aspekte gestalterischen Handelns eine stärkere Differenzierung des Rollenkonzepts und seiner Charakteristika vorgenommen. Diese differenzierte Perspektive zum Gestaltungshandeln dient in CTA ferner dazu, den daraus abgeleiteten Standpunkt für die Notwendigkeit eines „konstruktiven Assessments“ (Verbeek, 2006) von Gestaltungsprozessen zu begründen.

## Designer

Die Designer bezeichnen ein Rollenkonzept, das in der VSD-Literatur sehr prominent verwendet wird. Dabei kann zunächst der Eindruck entstehen, dass mit Designern primär visuelle oder konzeptuelle Gestaltungsfähigkeiten assoziiert werden, jedoch werden Designer in VSD häufig mit den technikschaftenden Developern und Engineers gleichgesetzt:

*„designers as far more important professionals since they not only can provide us with technical means but can also address the values of people and society and think about expressing them in material culture and technology“* (van den Hoven et al., 2015, S. 3).

Designer werden wie auch Entwickler und Ingenieure in ihrer Doppelfunktion als Gestalter auf technischer wie auch auf moralischer Ebene angesehen. Deshalb gilt auch für Designer, dass ihre Handlungen und Wertvorstellungen in die Betrachtung der Wertefragen betreffenden Aspekte in Gestaltungszusammenhängen miteinbezogen werden müssen und dass sie zudem für ihre getroffenen Entscheidungen verantwortlich gehalten werden. Wie bei den Entwicklern und Ingenieuren steht bei den Designern primär die Analyse der moralischen Dimension gestalterischen Handelns im Vordergrund, die vor der Zielsetzung eines reflektierten Gestaltungsprozesses und in einem Bewusstsein der damit verbundenen Herausforderungen durchgeführt wird. In VSD wird danach gestrebt, Designer zur Explikation und Vergegenwärtigung ihrer Annahmen und Wertvorstellungen zu motivieren und sie dabei zu unterstützen, um sie wiederum dazu zu befähigen, ihrer Verantwortung für ihr Gestaltungshandeln angemessen gerecht zu werden:

*„This is not to deny ignorance and uncertainty but rather requires designers to be more explicit about what they know and do not know and perhaps cannot know and to be more explicit about the value commitments they build into their designs so that they can assume accountability for their choices“* (van den Hoven et al., 2015, S. 5).

In VSD wird die Annahme vertreten, dass die Wertvorstellungen von Designern stark vom jeweiligen Gestaltungssetting und Arbeitskontext abhängen (Kontextsensitivität). Dies wird beispielsweise bei Friedman et al. (2013) deutlich, wo die Unterschiede zwischen in Unternehmen tätigen Designern und in Behörden tätigen Designern dargestellt werden. Für Designer in Unternehmen stellen häufig Umsatzsteigerung sowie Mitarbeiter- oder Kundenzufriedenheit Werte von zentraler Bedeutung dar, während für Designer, die in Regierungsbehörden aktiv sind, mehr nationale oder spezifische kommunale Werte im Vordergrund stehen, und sie zu dem Ziel beitragen, dass eine Behörde ihrem gesellschaftlichen Mandat gerecht werden kann. Hieran wird erneut deutlich, dass in VSD vertieftes Wissen über die konkreten Lebens- und Arbeitssituationen von Designern vorhanden ist, und dass in VSD zudem darauf geachtet wird, zwischen dem Selbstanspruch und dem Fremdanspruch von Designern zu differenzieren. Der Fremdanspruch kommt beispielsweise bei Batya Friedman zum Ausdruck, wenn sie beschreibt:

*„designers should provide users the greatest possible control over computing power“* (B. Friedman, 1996, S. 18).

Der Eigenanspruch von Designern zeichnet sich hingegen von Fragen des Typs „How do I translate stakeholder values into technical design decisions?“ (B. Friedman et al., 2017, S. 65) aus. Die in VSD vorgeschlagenen Methoden sollen Designer dazu befähigen, sich mit Wertefragen und ihren Implikationen produktiv und zielführend auseinanderzusetzen und für sich selbst sowie das zu gestaltende System wertvolle Erkenntnisse zu gewinnen (Borning & Muller, 2012). In VSD wird ferner die Annahme vertreten, dass ein deutlicher Mehrwert für die Designer selbst und für die Gestaltungsprozesse im größeren darin besteht, explizit transparent zu Wertefragen zu sein. In diesem Zusammenhang greift ebenso wieder der Standpunkt des progress, not perfection. Friedman und Hendry beschreiben die Intention hinter diesem Standpunkt und seiner Verankerung in VSD vor dem Hintergrund, die wahrgenommenen Barrieren von Designern bzgl. der Auseinandersetzung mit Wertefragen zu vermindern, sie mit den sprachlichen Werkzeugen auszustatten, die für derlei Auseinandersetzungen benötigt werden, und dadurch insgesamt möglicherweise vorhandene Bedenken und Vorbehalte auszuräumen:

*„it moves designers toward considering human values as a design criterion“* (B. Friedman & Hendry, 2019, S. 51).

Somit dient das Rollenkonzept in VSD auch dazu, Designer zur Reflexion zu ermutigen.

Auch in CTA wird das Rollenkonzept der Designer stellenweise dazu verwendet, um hervorzuheben und zu analysieren, inwieweit Handlungen von Gestaltern Werteaspekte berühren. Dies wird beispielsweise in der Arbeit von Peter-Paul Verbeek (2006) deutlich, der das Rollenkonzept dazu verwendet, um sich mit dem Aspekt der Wertemediation durch Technik zu befassen. Er zeigt auf, wie Designer durch ihre Arbeit stellenweise bewusst, stellenweise unbewusst moralische Entscheidungen treffen. Ihm dient die konkrete Vorstellung von Produktdesignern dazu, seine Gedanken darzulegen und zu erläutern, wie diese Berufsgruppe technische Bauteile und Systemkomponenten entwirft. Sie antizipieren wie Nutzer mit einem Produkt interagieren werden und geben implizit oder explizit bestimmte Nutzungspfade und Nutzungsmöglichkeiten für ihre Produkte vor. Damit delegieren sie bestimmte Verantwortlichkeiten an technische Artefakte. Sie schreiben Technik aber auch durch ihre Handlungen Werte ein (inscription). Der Autor beschreibt die Schwierigkeit, die für Produktdesigner in der Regel darin besteht, einzuschätzen, wie sich der Gestaltungskontext vom Nutzungskontext unterscheidet. Zur Bewältigung dieser Aufgabe kann nach seiner Einschätzung die CTA-Methodologie genutzt werden, da hier nicht nur die Designer eine Technologie evaluieren, sondern auch die künftigen Nutzer.

### 5.4.5 Die intervenierenden Prozessbegleiter

Die intervenierenden Prozessbegleiter bezeichnen das für CTA charakteristische Rollenkonzept der CTA agents. Sie zeichnen sich primär dadurch aus, dass sie Gestaltungsprozesse über lange Zeitverläufe hinweg begleiten, relevante Akteure identifizieren und die Gestaltungsgeschehnisse verfolgen, jedoch zu ausgewählten Zeitpunkten und mit gewählten Maßnahmen in das Gestaltungsgeschehen eingreifen. Sie werden in CTA als eine Art Zaungast konzipiert, die trotzdem, dass sie über Geschehnisse, Abhängigkeiten und Zukunftsszenarien von Gestaltungsprozessen im Bilde sind, eine gewisse Distanz zum inhaltlichen Gestaltungsgeschehen wahren. In CTA wird mit ihnen in Anspruch genommen, an zeitlich entscheidenden Stellen in Gestaltungsprozesse zu intervenieren und somit Einfluss auf das weitere Gestaltungsgeschehen zu nehmen. Allerdings wird mit CTA agents im Gegensatz zu den Rollenkonzepten der Gestalter *nicht* in Anspruch genommen, technische Systeme auf technischer oder anderweitig inhaltlicher Ebene zu gestalten. Das höchste Maß an Gestaltungsaktivität ist CTA agents dadurch zugebracht, dass sie in ihrer Funktion als Durchführende der CTA-Methoden (Kuhlmann, 2012) die Gestalter und andere Entscheidungsträger versammeln und deren Einstellungen und künftigen Entscheidungen zu beeinflussen suchen. CTA agents verhalten sich in der Vorstellung in CTA wie Change Agents (siehe Kap. 5.5.5), die als Experten für die Klärung von Konfliktsituationen und für die Steuerung von Innovationsprozessen Veränderungen in Gestaltungsprozessen herbeiführen:

*„CTA agents are change agents, but softly, through support and attempts at opening up, rather than pushing. If there is pushing, it is a push for more reflexivity“* (Rip & Robinson, 2013, S. 47).

Als Prozessbeobachter und Prozessbegleiter verfügen CTA agents über die notwendigen Kenntnisse zu den Fragen und Herausforderungen eines Gestaltungsprozesses, und nutzen dieses Wissen um unterstützende, perspektiven- und gedankenöffnende Maßnahmen durchzuführen, die in Gestaltungsprozessen bei den dort aktiven Gestaltern und Entscheidungsträgern ein stärkeres Bewusstsein für gesellschaftliche Interessen und mögliche Technikfolgen bewirken sollen. Sie besitzen die Aufgabe, die dafür notwendigen Veränderungs- und Reflexionsprozesse durch sanfte Interventionsmaßnahmen (soft interventions) anzustoßen. Eine Besonderheit, die sie im Vergleich mit den anderen im Gestaltungsprozess aktiven Rollen auszeichnet, ist ihre Fähigkeit sich zwischen verschiedenen Orten und Gruppen im Technologiefeld zu bewegen und somit Wissen über die dort vorherrschenden Gesamtzusammenhänge herzustellen. Diese räumliche Mobilität ist essentiell für CTA agents, damit sie ihre Arbeit zielführend verrichten können. Sie erfordert von ihnen selbst jedoch auch eine hohe gedankliche Flexibilität und die Fähigkeit, sich immer wieder an dynamische Gestaltungssettings anzunähern und anzupassen. Durch ihre Positionierung außerhalb des Kreises von Entscheidungsträgern sind sie zudem nicht weisungsgebunden und sind üblicherweise weniger direkt in den Fortschritt oder den Ausgang eines Gestaltungsprozesses investiert.

Um ihren Aufgaben nachzukommen, verfügen CTA agents über Expertise im Umgang und in der Anwendung sozialwissenschaftlicher Methoden (Rip & Robinson, 2013, S. 47). Dieses Qualifikationsprofil wird in CTA besonders für die Durchführung der Technologiefeldanalysen als kritisch betrachtet, da nur dort das Wissen erarbeitet werden kann, auf dessen Grundlage sämtliche weitere Aktivitäten der CTA agents aufbauen. Auch benötigen sie diese Expertise, um knifflige Entscheidungen zur methodischen Umsetzung in CTA zu treffen:

*„For the broader aims of Constructive TA, there is an uneasy trade-off between CTA agents accommodating to existing enactors' perspectives (so as to keep the enactors involved) and introducing incentives to broaden their perspectives (so as to induce some change)“ (Rip & te Kolve, 2008, S. 56).*

Spätestens an dieser Stelle wird erkennbar, dass das Rollenkonzept der CTA agents in CTA die zusätzliche Funktion besitzt, ein Selbstportrait der Gruppen und Individuen darzustellen, die TA-Arbeit verrichten. Die CTA-Literatur führt diese Funktion nicht explizit auf, macht dies jedoch deutlich, indem sie CTA agents als die durchführenden Personen der CTA-Methoden positioniert. Das Rollenkonzept dient den Autoren, die das Rollenkonzept verwenden, einerseits zur Reflexion eines Status Quo von TA-Arbeit sowie der Darlegung von Gedanken zu künftiger TA-Arbeit. Konkret wird mit dem Rollenkonzept der CTA agents der Versuch unternommen, darzulegen, wie sich die Arbeit und das Selbstverständnis von TA-Arbeit verrichtenden Personen weiterentwickeln sollten, um einen bedeutsamen Beitrag zu einer sozialverträglichen TG leisten zu können. Somit besitzt das Rollenkonzept der CTA agents eine für die TA-Community nach innen gerichtete Reflexionsfunktion, die die Notwendigkeit der Adaption bzw. Weiterentwicklung bestimmter traditioneller TA-Positionen fordert. Durch diese selbstreferentielle Funktion wird auch erkennbar, weshalb die CTA-Literatur trotz einem gewissen Anspruch an direkter Interventionsaktivität und damit verbundener indirekter Auswirkung auf Gestaltungsvorgänge stets die Positionierung der CTA agents außerhalb der direkten Entscheidungsträger hervorhebt – dies ist die Position in der sich TA-Arbeit verrichtende Gruppen und Individuen traditioneller Weise positionieren und in der sie sich nach ihrer Einschätzung befinden müssen, um den an sie gestellten Anforderungen auf angemessene Weise gerecht zu werden.

Mit der selbstreferentiellen Funktion des Rollenkonzepts geht ein weiterer Aspekt einher, der eine mehr übergreifende Relevanz in CTA besitzt. Die Durchführung der Technologiefeldanalysen und CTA-Workshops bringt die Notwendigkeit mit sich, dass CTA agents sich aktiv in ein Feld begeben und durch das Hineinbegeben in ein Feld bereits Veränderungen herbeiführen, die ohne eine solches Hineinbegeben (insertion) nicht zustande kämen:

*„visiting “knowledgeable” strangers irritate existing ways of working and thus create openings for learning and further evolution of how we handle new technologies in our society“ (Rip & Robinson, 2013, S. 51).*

Dies ist ein klassisches Dilemma qualitativer Forschungsarbeit und somit auch in CTA relevant. Die CTA-Literatur versucht diesem Dilemma mit Offenheit zu begegnen, indem es einerseits Argumente für die Notwendigkeit einer Insertion von CTA agents in Technologiefelder aufzeigt und damit verbunden die Unausweichlichkeit von Veränderungen hervorhebt. Anders als andere Positionen in TA versucht CTA durch das Rollenkonzept der CTA agents jedoch, diese Veränderungen nicht unter großem Aufwand zu verhindern, sondern den Umstand zu akzeptieren und mit ihm möglichst produktiv zu arbeiten. Durch diese Position und Einschätzung fungiert das Rollenkonzept somit auch als Argumentationshilfe für die zuvor angesprochene, als Notwendigkeit erachtete Adaption einiger traditionellen Positionen in der TA-Community. Besonders in Texten von Arie Rip wird diese Adaption deutlich gefordert, weshalb das Rollenkonzept der CTA agents durchaus auch einen normativen Charakter besitzen kann.

Darüber hinaus werden in CTA weitere reflektierende Gedanken im Zusammenhang mit dem Rollenkonzept der CTA agents angestoßen. Für CTA agents gilt, wie für die Rollenkonzepte der Gestalter, dass mit ihrer Arbeit aufgrund der damit verbundenen Intention, Veränderungen in Gestaltungsprozessen herbeizuführen, eine gewisse moralische Verantwortung verbunden ist. Die CTA-Literatur weist darauf hin, dass CTA agents sich dieser Verantwortung bewusst sein sollten und dass sie ihr durch ihr Verhalten auf angemessene Weise Rechnung tragen müssen:

*„CTA agents have a moral thrust and moral justification of their being active as a CTA agent. Thus, they are also moral entrepreneurs, but on a meta level“*  
(Schot & Rip, 1997, S. 265).

Dies wird in der Literatur ferner zum Anlass genommen, das Werteprofil von CTA agents zu diskutieren: „CTA agents embrace values such as being anticipatory, reflexive, and oriented toward learning“ (Schot & Rip, 1997, S. 265). Diese Werte repräsentieren zentrale Charakteristika für CTA agents und besitzen in CTA deshalb einen durchgängig hohen Stellenwert, wenn CTA agents referenziert werden. Für CTA agents als Anwender der CTA-Methoden gilt es darüber hinaus, ihre Werte zum Bestandteil ihrer eigenen Artikulationsstrategie zu machen, um ihre eigenen Positionen und Überzeugungen darlegen zu können. Dies soll es wiederum den Technologiefeldakteuren erleichtern, die Notwendigkeit für eine stärkere an gesellschaftlichen Fragestellungen orientierte TG zu erkennen und im Idealfall in ihre eigene Artikulationsstrategie zu übernehmen.



### 5.4.6 Die Nutzer

Das Rollenkonzept der Nutzer ist in CTA und VSD das typische Gegenstück-Rollenkonzept zu den Rollenkonzepten der Gestalter. Nutzer repräsentieren Gruppen und Individuen, die in der Regel über eine geringere technische Expertise wie Gestalter verfügen, und deren primäres Interesse sich an die Nutzung eines technischen Systems richtet, um bestimmte Aufgaben in ihrem Privat- oder Berufsleben durchführen zu können. Sie repräsentieren primäre und deshalb häufig offensichtliche Zielgruppen für die ein technisches System gestaltet wird, wodurch ihre Relevanz für einen Gestaltungsprozess gekennzeichnet ist. In der Literatur zu beiden Ansätzen wird darauf hingewiesen, dass das Ermitteln und Berücksichtigen konkreter Attribute von Nutzergruppen von zentraler Bedeutung ist (Kontextsensitivität), weshalb in den mehr übergreifenden Ausführungen zu beiden Ansätzen mehr allgemeine Aussagen zum Rollenkonzept getroffen werden. Somit ist auch das Rollenkonzept der Nutzer häufig ein Platzhalter-Rollenkonzept in vielen CTA- und VSD-Arbeiten, deren konkrete Ausprägungen sich je nach betrachtetem Gestaltungskontext unterscheidet und über die entsprechend wenig Charakteristika kommuniziert werden, deren Gültigkeitsanspruch sich auf sämtliche CTA- und VSD-Arbeiten übertragen lässt. Ein Beispiel für eine konkrete Ausprägung des Rollenkonzepts in VSD findet sich bei Yoo et al. (2013), in deren Studie Einsatzszenarien von Mobiltelefonen zur Schaffung von Sicherheit exploriert werden. Zu den in diesem Kontext identifizierten konkreten Ausprägungen des Nutzer-Rollenkonzepts zählen die Gruppen obdachloser jungen Menschen, Service-Provider von Mobilfunkdiensten sowie Angehörige des Polizeiberufs.

In der VSD-Literatur wird das Rollenkonzept der Nutzer zudem häufig mit den Stakeholdern, konkret mit den direct Stakeholders gleichgesetzt, die mit einem technischen System interagieren. Wie bereits bei den Rollenkonzepten der Gestalter dient in VSD somit die Interaktionsebene als Hilfsmittel bzw. Ansatzpunkt, um Nutzer zu identifizieren und zu beschreiben (interactional stance). Durch die Verwendung als Synonym zu direkten Stakeholdern ist die Funktion des Rollenkonzepts somit eng verknüpft mit den zuvor erläuterten Funktionen des Rollenkonzepts der Stakeholder (siehe Kap. 5.4.1). In VSD dient das Rollenkonzept zudem primär dazu, die Werte von Nutzern zu beschreiben und zu untersuchen sowie nachvollziehbar darzulegen, weshalb die Ermittlung und Auseinandersetzung mit den Werten von Nutzern in Gestaltungsprozessen notwendig und zielführend ist. Die Ergründung und Formulierung von Nutzerwerten wird in vielen VSD-Arbeiten als ein zentraler Baustein von Gestaltungsprozessen erachtet, anhand deren Inhalte und Bedeutungen wiederum relevante Systemeigenschaften abgeleitet oder definiert werden, um sie durch gestalterische Aktivitäten zu adressieren. Dabei wird sich häufig einer grundlegenden, für Gestaltungsprozesse traditionellen Vorstellung eines Endverbrauchers (end users) eines technischen Produkts oder Service bedient, besonders in VSD wird diese Vorstellung jedoch bewusst um die explizite Nennung der indirect stakeholders erweitert. In VSD werden viele Anstrengungen unternommen, um sich durch dieses deutlich erweiterte Nutzerverständnis von anderen Gestaltungsansätzen abzugrenzen und zu verdeutlichen, welche Vorkehrungen zur Berücksichtigung von Nutzern und ihren Werten getroffen werden. In VSD kommt ein ganzheitliches Nutzerverständnis

zum Tragen, das den Blick bewusst auf diejenigen Nutzergruppen richtet, die nicht direkt mit einem System interagieren können:

*„The intent was to develop broader language and, thereby, to create space to reach beyond only users to consider systematically those affected by a given technology“* (B. Friedman & Hendry, 2019, S. 40).

Die mit dem Rollenkonzept verbundenen Vorstellungen reichen weiter wie die vergleichbarer Gestaltungsansätze, in welchen die mehr ökonomisch angesetzte Betrachtung von Nutzern als Verbraucher oder Konsumenten für die Betrachtung von Wertefragen betreffenden Aspekten zu unterkomplex sind. Aus diesem Grund wird in VSD für die Notwendigkeit einer erweiterten und ganzheitlichen Perspektive auf Nutzer argumentiert. Unabhängig von dieser Erweiterung kommen auch in VSD viele der mehr traditionellen Verwendungszwecke des Rollenkonzepts zur Anwendung. So dient es beispielsweise häufig dazu, zwischen verschiedenen Aufgaben der am Gestaltungsprozess Beteiligten zu unterscheiden, um beispielsweise das Rollen-Rechte-Konzept eines IT-Systems festzulegen. Wie bei den anderen Rollenkonzepten in VSD steht auch bei den Nutzern häufig der Aspekt der Wertkonflikte im Vordergrund. So wird beispielsweise in der Arbeit von Dodier-Lazaro et al. (2017) deutlich, dass im Zusammenhang der Gestaltung mit Systemen, die Nutzern Sicherheit gewähren sollen (z.B. Passwort-Manager), häufig in Konflikt mit der Usability selber Systeme stehen:

*„Usable security experts continue to disregard the preferences and priorities of users, the economic implications of deploying non-user-centred security, and the disconnect between how security is valued by security experts and valued by users. This lack of attention to users’ values has led to unusable, ineffective and unused security mechanisms“* (Dodier-Lazaro et al., 2017, S. 5).

Im Vergleich zu VSD findet das Rollenkonzept der Nutzer in der CTA-Literatur deutlich seltener Verwendung. Auch hier wird es wenn, dann häufig als Synonym zu den Stakeholdern verwendet. So definieren Schot und Rip an einer Stelle, dass die Bedürfnisse von Nutzern einen Faktor von Gestaltungsprozessen darstellen, der bei der Durchführung von CTA-Aktivitäten Aufmerksamkeit geschenkt werden muss, besonders um das viel referenzierte societal learning zu ermöglichen:

*„Learning must occur, and it should be broad and deep. Broad learning entails an exploration of possible new linkages between a range of aspects such as design options, user demands, and issues of political and societal acceptability“* (Schot & Rip, 1997, S. 257).

Sowohl in CTA wie auch in VSD werden Nutzer interpretiert als Gruppen und Individuen, die mit einem technischen System interagieren. Jedoch dient diese Interpretation in beiden Ansätzen häufig auch als Ausgangspunkt, um die mit dem Rollenkonzept verbundenen Defizite und Beschränkungen aufzuzeigen, die mit seiner traditionellen Verwendungstypischerweise einhergehen. In VSD wird hierfür die Bedeutung von Nutzerwerten und Wertkonflikten zugrunde gelegt, während in CTA die noch weitere Betrachtung auf gesellschaftlicher Ebene Anwendung findet.

### 5.4.7 Die Technikenthusiasten

Die Technikenthusiasten bezeichnen Rollenkonzepte, die im Kontext von CTA eine besondere Bewandnis besitzen. In CTA werden sie als Gruppen und Individuen beschrieben, die eine besonders stark ausgeprägte Überzeugung von Technik im Allgemeinen oder von ausgewählten technischen Systemen oder Anwendungen besitzen und es als ihre Aufgabe betrachten, ihre Mitmenschen von ihrer Sichtweise zu überzeugen. Dabei kann es sich sowohl um einen Aufgabenbestandteil eines bestimmten Berufsprofils handeln, es kann jedoch auch die Verrichtung einer Arbeit aus persönlicher Überzeugung beschreiben. Zu den Rollenkonzepten der Technikenthusiasten in CTA zählen die Promoters und die Proponents. Auch wenn in VSD der Aspekt der Begeisterung für Technik stellenweise von Bedeutung ist, verwendet die VSD-Literatur in der Regel keine speziell dafür entwickelten Rollenkonzepte.

#### Promoters & Proponents

Die typische Repräsentation von Technikenthusiasten in CTA stellt das Rollenkonzept der Promoters da. Sie sind Befürworter einer bestimmten Technologie oder einer technischen Anwendung, und fördern deren Gestaltung, Weiterentwicklung oder Einsatz. Sie werden in der CTA-Literatur stellenweise als Synonym zu den Enactors (siehe Kap. 5.4.2) verwendet, da sie es gemäß der Interpretation in CTA ähnlich wie die Enactors als ihre Aufgabe sehen, mögliche Hürden auf dem Weg zum gesellschaftlichen Einsatz von Technologie aus dem Weg zu räumen. Stellenweise werden Promoters auch mit Rollenkonzepten der Gestalter gleichgesetzt, da sie häufig dieselben Charakteristika aufweisen: „Enactors are technology developers and promoters aiming to enact new technology“ (Kuhlmann, 2012, S. 3). Somit muss in CTA nicht immer zwingend ein Unterschied zwischen Enactors, Developers und Promoters bestehen, jedoch wird an mehreren Stellen der CTA-Literatur die für Technik werbende Aktivität der Promoters explizit in den Vordergrund gestellt. Sie ist in CTA von besonderer Relevanz, weil Promoters im Zuge ihrer Arbeit häufig Wissenschaftskommunikation betreiben, die die vermeintlichen Vorteile einer Technologie teilweise überformt wiedergibt, während mögliche negative Konsequenzen nicht behandelt und verschwiegen werden. Stellenweise wird in der CTA-Literatur auch angedeutet, dass sich die konkreten Maßnahmen der Promoters auch in Form von Marketingkampagnen äußern. Durch ihre kommunikative Funktion erfahren Promoters häufig auch außerhalb von Fachkreisen viel Aufmerksamkeit, wodurch sie in CTA auch in Betrachtungen von Verhaltensweisen und Argumentationsstrategien der

societal actors relevant werden. Diese stützen sich häufig auf von Promoters kommunizierten Positionen, die es in CTA-Projekten gilt zu reflektieren und mögliche vorhandene inhaltliche Lücken und argumentative Fallstricke in den Vordergrund zu rücken:

*„Thus, promotion actors (engineers and others) need not realize that when they are engineering technology they are also engineering society. Many of the negative or undesirable outcomes of the activities of the promoters derive from this lack of insight“* (Schot & Rip, 1997, S. 264).

Somit stellt das Rollenkonzept der Promoters eine bewusst durchgeführte Präzisierung von Gestalter-Rollenkonzepten dar, die die werbende Facette dieser Rollenkonzepte in den Vordergrund rückt, um bewusst Kritik an dieser zu üben. Da die Zielgruppe der Promoters technisch weniger versierte Gruppen und Individuen darstellen, können diese häufig den Wahrheitsgehalt der von Promoters getätigten Aussagen nur schwer verifizieren. In CTA wird deshalb bewusst auf die Verantwortung hingewiesen, die Promoters mit der Durchführung ihrer werbenden Aktivitäten besitzen. Diese Verantwortung gilt es nach Ansicht in CTA wahrzunehmen und sich nicht (wie bei den Enactors) möglicherweise berechtigten Zweifeln zur Aufgabe zu machen, diese schlicht zu entkräften.

Die Proponents sind im Grunde ein Synonym zum Rollenkonzept der Promoters, allerdings kommt bei Ihnen auf sprachlicher Ebene noch deutlicher ihre Gegenposition zu den Rollenkonzepten der Technikgegner, speziell der Opponents (siehe Kap. 5.4.8) zum Ausdruck. Bei der Wortwahl der Proponents wird zudem bewusst mit der Mehrdeutigkeit des Begriffs gespielt, den Promoters oder Proponents können im ökonomischen Sprachgebrauch auch Gruppen oder Individuen bezeichnen, die nicht nur ideell von einer Technologie überzeugt sind, sondern auch faktisch für deren Gestaltung, Verbreitung und Nutzung innerhalb einer Organisation verantwortlich sind. Dieser häufig ökonomisch motivierte Hintergrund dient in CTA als Anlass, zu begründen, weshalb in Gestaltungsprozessen mehr Reflexivität benötigt wird. Denn Proponents und Promoters fehlt nach Ansicht in CTA diese Reflexionsfähigkeit:

*„Reflexivity is also needed, and for two reasons. First, reflexivity about technology is necessary to avoid falling back in a naive contrast between technology and society, and its attendant proponents-opponents dichotomies try to reduce space for negotiating to the direction and nature of technical change“* (Schot & Rip, 1997, S. 257).

Beiden Rollenkonzepten wird eine häufig wenig hinterfragte Begeisterung von Technik zugeschrieben, die im äußersten Fall dazu führen kann, dass mit Technik verbundene Risiken und Gefahren bewusst oder unbewusst von Proponents und Promoters verharmlost werden. Hierin identifiziert CTA ein Potential für eine Fehlinformation gesellschaftlicher Akteure, der es durch die Anwendung von CTA-Methoden vorzubeugen gilt. Somit dienen beide Rollenkonzepte in CTA auch dazu, den Grad der Naivität von Gestaltern zu benennen, sowie zu untersuchen, inwieweit sie bei der Durchführung ihrer Aktivitäten vorsätzlich oder versehentlich mit Intransparenz arbeiten.

#### 5.4.8 Die Technikgegner

Die Technikgegner bezeichnen das Gegenstück-Rollenkonzept zu den Technikenthusiasten in CTA, die Opponents. Sie zeichnen sich primär dadurch aus, dass sie die Begeisterung und Überzeugung der Technikenthusiasten für eine bestimmte Technologie und deren Anwendung weniger teilen und entsprechende Gegenpositionen vertreten. Sie stellen einen Gegenpol zu Positionen der Technikenthusiasten dar, deren Hintergründe es in CTA-Projekten zu ergründen gilt. Wie bei den Technikenthusiasten gilt auch für die Opponents, dass die mit ihnen verbundenen Positionen und Argumente grundsätzlich auch in VSD thematisiert werden, wo hauptsächlich im Kontext von Wertekonflikten Befürworter und Gegner von beispielsweise Gestaltungsmaßnahmen oder Systemfeatures vorhanden sind. Jedoch wird in VSD für die Darstellung solcher Zusammenhänge und Konflikte kein eigens und explizit formuliertes Rollenkonzept verwendet, weshalb sich die folgenden Ausführungen primär auf die CTA-Literatur beziehen.

Die grundsätzliche Prämisse der Opponents ist in CTA recht simpel nachzuvollziehen – sie sind gegen ein bestimmtes technisches Feature, die Realisierung eines Anwendungsfalls oder gar der Nutzung einer Technologie als solches. Sie bekleiden die Contra-Seite einer Pro-Contra-Diskussion. Durch die Darstellung von Gestaltungsprozessen als Ausfechtungsprozesse zwischen Opponents und Proponents wird erkennbar, dass Gestaltungsprozesse in CTA durchaus auch als politische Aushandlungsprozesse betrachtet werden. Wie auch bei politischen Debatten werden die Aushandlungen zwischen Opponents und Proponents in CTA als eine Art Wettstreit sämtlicher Pro- und Contra-Argumente konzipiert, die von den verschiedenen beteiligten Akteuren vorgetragen und angehört werden. Auf Basis der auf diese Weise zustande kommenden Informationslage kann anschließend eine Konsensfindung über die Gestaltung oder den weiteren Verlauf des Einsatzes einer Technologie entschieden bzw. abgestimmt werden. Diese Perspektive schließt an das ko-evolutionäre Verständnis von Gestaltungsprozessen in CTA sowie an die Darstellung von Machtgefügen und Machtgefällen im Zusammenhang mit den Rollenkonzepten der informierten Einflussnehmer (siehe Kap. 5.4.2) an. Allerdings bezeichnet das Rollenkonzept der Opponents auch eine Besonderheit in CTA, denn im Vergleich zu VSD wird in der CTA-Literatur mehrfach explizit darauf hingewiesen, dass in sämtlichen Gestaltungskontexten *immer* Opponents vorzufinden sind. Somit geht auch mit diesem Rollenkonzept ein übergreifender Gültigkeitsanspruch einher, der sich unabhängig einer betrachteten Technologie auf sämtliche Gestaltungsprozesse übertragen lässt. Dieser Anspruch ermöglicht es in CTA, anhand der von Opponents vertreten Positionen in Gestal-

tungsprozessen bestimmte Grundsatzfragen zu stellen oder bestimmte Grundsatzargumente zur Diskussion zu stellen, die möglicherweise von den Technologiefeldakteuren bereits als gegeben angenommen werden und deshalb ihrer Ansicht nach keiner weiteren Aufmerksamkeit bedürfen. Auch dient das Rollenkonzept der Opponents dazu, auf die Verlierer von Gestaltungsprozessen bzw. die Rollenkonzepte der Unbeteiligten und Betroffenen hinzuweisen. Diese vertreten typischerweise auch Gegenargumente für bestimmte Entwicklungen oder Einsatzzwecke von Technik, deren Gehalt nach Ansicht in CTA deutlich mehr Aufmerksamkeit gewidmet werden sollte. Ihre Positionen werden häufig auf gleiche Weise als berechtigt und valide betrachtet wie jene der Technikenthusiasten. Dadurch dient das Rollenkonzept in CTA auch dazu, eine Ungleichbehandlung der Positionen von Opponents und Proponents festzustellen, wodurch in CTA wiederum eine Notwendigkeit für die in ihm vorgeschlagenen Interventionsmaßnahmen abgeleitet wird.

Eine Aufteilung sozialer Prozesse in Befürworter und Gegner kann wie zuvor auf die Metapher einer politischen Debatte übertragen werden, jedoch geht damit noch eine weitere Funktion einher. Die Aufteilung eines komplexen Ausschnitts sozialer Realität (Gestaltungsprozesse) in Befürworter und Gegner lässt vermeintlich vorhandene Grenzen zum Vorschein treten, die es den involvierten Gruppen und Individuen ermöglicht sich zu positionieren und eine Umgangsform und Manövrierfähigkeit für die vorgefundene Komplexität zu finden. Somit kommt beim Rollenkonzept der Opponents auf besondere Weise seine komplexitätsreduzierende Funktion zum Vorschein. Dies wird an einer Stelle der CTA-Literatur auch explizit zum Ausdruck gebracht:

*„Neither 'proponents' nor 'opponents' are simple categories, but to think in those terms seems natural. And the labels are used prospectively: introducers of new technology now expect that there may be contestation, and watch out for 'opponents.' The opponent/proponent dichotomy has become a pattern in our culture, and it serves actors in their attempts to order a complex environment“*  
(Rip, 2001, S. 111).

## 5.5 Ursprünge der Rollenkonzepte

Durch die verschiedenen Bedeutungen und Verwendungszwecke der Rollenkonzepte wird ersichtlich, welche vielfältigen Annahmen zu menschlichem Leben und Handeln in CTA und VSD vorhanden sind. Die Rollenkonzepte sind zentraler Bestandteil einiger bedeutsamer Argumentationslinien in den beiden Ansätzen und sind ferner maßgebliche Instrumente in der Ausarbeitung und letztlich Anwendung der in den Ansätzen vorgeschlagenen Methoden zur Gestaltung technischer Systeme. Die Verwendung einiger Rollenkonzepte ist dabei recht naheliegend, da es sich um Rollenkonzepte handelt, deren Verwendung in Kontexten zur TG recht üblich sind. Dies zeigen die Verwendungsweisen der Rollenkonzepte der Gestalter, Nutzer, Involvierten und Unbeteiligten. Ihre Interpretationen weisen zudem Anknüpfungspunkte zu Positionen und Argumentationslinien der MZG auf. Bei manchen der Rollenkonzepte handelt es sich jedoch um für die Ansätze

charakteristische, sowie stark implizite gedankliche Konstrukte, deren Bedeutungsermittlung schnell an Grenzen stößt, wenn die Ursprünge der Rollenkonzepte außer Acht gelassen werden. Dies liegt darin begründet, dass die Ausarbeitung und Verwendung mancher Rollenkonzepte stellenweise auf Wissensbestände und Gedankengänge zurückgreift, die der Entwicklung von CTA und VSD zeitlich vorausgegangen sind und die stellenweise auch in anderen Fachbereichen erarbeitet wurden. Die Gründe für Übernahmen und Adaptionen von Rollenkonzepten variieren mit den mit ihnen intendierten Zwecken in VSD und CTA, weshalb in ihrer Betrachtung ein besonderer Informationsgehalt für die Fragestellung der Dissertation steckt. Einblicke in die Hintergründe solcher Übernahmen und Adaptionen versprechen vertiefende Erkenntnisse bezüglich der in den Ansätzen zum Tragen kommenden Interpretationen der Mensch- und Werteorientierung. Meine nachfolgenden Erläuterungen widmen sich daher verstärkt diesen vertiefenden Hinweisen, die sich in der Literatur auf die Ursprünge der Rollenkonzepte finden lassen. Sie vervollständigen das bereits ausgearbeitete Verständnis der Rollenkonzepte und geben vertiefende Einblicke in die Denk- und Arbeitsweisen, die in beiden Ansätzen Anwendung finden.

### 5.5.1 Die Involvierten

Die Involvierten bezeichnen die „Platzhalter“-Rollenkonzepte der Stakeholder und Akteure, die sich in beiden Ansätzen typischerweise durch sehr wenig definierte Attribute und einen umfassenden Gültigkeitsraum auszeichnen. Zu den Hintergründen der Unterscheidung direkter und indirekter Stakeholder finden sich in der VSD-Literatur vertiefende Hinweise, während sich zum Rollenkonzept der Actors in der CTA-Literatur vertiefende Hinweise finden.

#### Stakeholder

Zur Verwendung des Rollenkonzepts der Stakeholder beschreiben Friedman und Hendry an einer Stelle, dass das Rollenkonzept in VSD Anwendung findet, um sämtliche Personen und weitere Entitäten beschreiben zu können, die von einer Technologie betroffen sein können:

*„Value sensitive design adopted “stakeholder,” a term with a long history, to refer to those people or entities who are or will be significantly implicated by a technology“ (B. Friedman & Hendry, 2019, S. 39).*

Hintergrund dieser Übernahme ist der Umstand, dass das Rollenkonzept der Nutzer in den 1990er-Jahren in den Communities verwendet wurde, aus denen VSD hervorging. In diesen Communities bezeichnet es jedoch häufig die direkt mit einem System interagierenden Menschen. In VSD wird nach Friedman und Hendry jedoch ein Rollenkonzept benötigt, das diese stark gerichtete Perspektive auf Mensch-Technik-Verhältnisse erweitert, um neben Nutzern als direkten Stakeholdern die besondere Bedeutung indirekt Betroffener zum Ausdruck bringen und untersuchen zu können. Das Stakeholderrollenkonzept stellt eine dafür notwendige, grundsätzlich höhere Abstraktionsebene zur Verfügung. In den erläuternden Aussagen der beiden Autoren wird zudem auf die Stakehol-

der Theory als Anwendungsgebiet des Rollenkonzepts verwiesen. Dort entleihen die beiden Autoren eine von Freeman (1984) formulierte Definition, die Stakeholder als sämtliche Gruppen oder Individuen beschreibt, die das Erreichen der Ziele einer Organisation beeinflussen können oder von diesen Zielen betroffen sein können (Mitchell et al., 1997, S. 856). Diese Definition findet häufig im Fachgebiet des Organizational Management Anwendung und erweitert das Stakeholder- bzw. Nutzerverständnis dahingehend, dass der betrachtete realweltliche Ausschnitt, in dem es verwendet wird, sich auf Organisationen bezieht. Jedoch ist auch diese Rahmung der Betrachtung und Verwendung des Rollenkonzepts für die Zwecke in VSD noch zu limitierend, da VSD auch zur Gestaltung von technischen Systemen in organisationsübergreifenden Strukturen oder gänzlich außerhalb von Organisationskontexten angewendet werden können soll. Friedman und Hendry stützen sich deshalb für ihre Auslegung des Rollenkonzepts mehr auf Arbeiten aus der HCI (Shneiderman & Rose, 1996) und des PD (M. J. Muller, 1995), in denen mit Stakeholdern auch Personen bezeichnet werden, die keiner Organisationseinheit zugeordnet werden können, jedoch dennoch von Auswirkungen technischer Systeme betroffen sein können. Diese Ausführungen von Friedman und Hendry legen nahe, dass mit der Bedeutung und Verwendung des Stakeholderrollenkonzepts in VSD auch eine Kritik der Eindimensionalität des in Gestaltungsprozessen etablierten Nutzer-Rollenkonzepts der 1990er- und frühen 2000er-Jahre verbunden ist.

Für die Formulierung ihrer um die indirekten Stakeholder erweiterten Definition des Rollenkonzepts verweisen Friedman und Hendry zudem auf eine Arbeit aus dem Fachgebiet des Environmental Management (Reed, 2008), die sich mit den Möglichkeiten und Herausforderungen der Partizipation von Stakeholdern in Gestaltungs- und Entscheidungsprozessen auseinandersetzt. Aus diesem Zusammenhang wird eine weitere Funktion des Rollenkonzepts in VSD ersichtlich, denn auch wenn in VSD die Einschätzung geteilt wird, dass die Beteiligung von betroffenen Stakeholdern an Entscheidungsprozessen sich als herausfordernd darstellen kann, wird Stakeholderpartizipation in VSD grundsätzlich befürwortet. Zudem wird aus der Referenz auf das die Arbeit von Reed ersichtlich, weshalb im Stakeholderverständnis in VSD die für Gestaltungsprozesse eher untypische Berücksichtigung nicht-menschlicher Entitäten erwähnt wird – Fragen des Umweltmanagements berühren häufig weitreichende, komplexe Zusammenhänge und beziehen sich häufig auf sämtliche Lebewesen auf der Erde, wodurch neben den Menschen vorwiegend auch Tiere und Pflanzen im Fokus der Untersuchungen liegen. Auch streben Umweltmanagementprozesse an, Entscheidungsfindungen zu unterstützen, deren Effekte sich auf sehr lange zukünftige Zeiträume auswirken können und von denen sich Entscheidungsträger langfristig positive Effekte für sämtliche Lebewesen der Erde und ihrer Natur versprechen. Diesen deutlich weiteren Horizont streben Friedman und Hendry an, in VSD aufzugreifen, auch wenn in VSD mit dem Rollenkonzept der Stakeholder nicht in Anspruch genommen wird, Prozesse oder Phänomene globalen Ausmaßes zu adressieren. Es geht ihnen vielmehr darum, den Betrachtungs- und Gedankenhorizont der in Gestaltungsprozessen beteiligten Gruppen und Individuen über die üblichen Grenzen hinweg zu erweitern, um die relevanten Aspekte der räumlichen, natürlichen und gedanklichen Umwelt eines Gestaltungsprozesses zu erkennen und in diesem zu berücksichtigen. Dies



wird auch dadurch verdeutlicht, dass mit Stakeholdern in VSD beispielsweise auch Gebäude (als Vertreter für von Menschenhand geschaffene Artefakte), zukünftig lebende Generationen oder auch „sacred mountaintops“ (B. Friedman & Hendry, 2019, S. 40) bezeichnet werden können. Diese Förderung alternativer und erweiterter Perspektiven auf das Gestaltungsgeschehen, der daran beteiligten und von ihm betroffenen Entitäten soll Anwender von VSD zu einem **Denken „outside the box“** motivieren und ihnen so die Bedeutungen und Tragweiten von Wertefragen berührenden Aspekten in Gestaltungsprozessen aufzeigen. Auf diese Weise trägt das Rollenkonzept der Stakeholder in VSD auch dazu bei, die an Gestaltungsprozessen beteiligten Gruppen und Individuen für größere Zusammenhänge oder möglicherweise relevante Details in Gestaltungsprozessen zu sensibilisieren, um so zu einer „wertsensitiven“ Gestaltung beizutragen.

Auch wenn in der CTA-Literatur das Rollenkonzept der Stakeholder weitgehend unkommentiert verwendet wird, wird es auch hier stellenweise mit seinem Ursprung in der Stakeholder Theory referenziert (Beekman, 2004; Braun & Starkbaum, 2023; Quist & Vergagt, 2000; Steen & Nauta, 2020). Wenn sich die CTA-Literatur mit den Hintergründen des Rollenkonzepts auseinandersetzt, erfolgt dies jedoch häufig zu ähnlichen Zwecken wie in VSD, so werden beispielsweise die Notwendigkeit von Stakeholderpartizipation in Gestaltungsprozessen, Best Practices für seine Realisierung sowie die Chancen und Risiken einer solchen diskutiert. Derartige Arbeiten sind daher instruktiv für das Nachvollziehen des Stakeholderverständnisses in CTA, gehen jedoch nicht über die bereits dargestellten Einblicke in die Bedeutung und Verwendung des Rollenkonzepts in VSD und CTA hinaus.

## Actors

Zum Rollenkonzept der Actors in CTA erläutert Konrad (2021, S. 211), dass dessen Verwendung vor dem Hintergrund geschieht, Akteure anhand ihrer verschiedenen sozialen Rollen unterscheiden zu können. Der Begriff der sozialen Rolle verweist auf die Soziologie, wo das Rollenkonzept der Actors beispielsweise auch Teil des Begriffssystems der Rollentheorie ist (siehe Kap 5.2). In der Soziologie kommt die Vorstellung von Akteuren als Inhaber sozialen Rollen typischerweise zum Einsatz, um Sozialisations- und gesellschaftliche Veränderungsprozesse zu untersuchen oder um das Aufkommen von Konflikten zwischen gesellschaftlichen Gruppen oder Inhabern multipler sozialer Rollen nachzuvollziehen. Derartige Verwendungsweisen kommen entsprechend auch in Subdisziplinen der Soziologie wie der Techniksoziologie zum Einsatz, sowie den von der Soziologie geprägten interdisziplinären Fachbereichen wie den Science and Technology Studies (STS). Durch die Fokussierung sozialer Prozesse in Kontexten von Technikaneignung und TG wird das Rollenkonzept der Actors in solchen Feldern jedoch besonders zu dem Zweck verwendet, Akteurskonstellationen zu beschreiben und ihre Ziele und Handlungsabsichten zu ergründen oder zu extrapolieren. Diese Zwecke erfüllt das Rollenkonzept auch in CTA, was naheliegt, da einige zentrale Positionen in CTA sowie deren Herleitung sich auf techniksoziologische und auch STS-Theorien wie das SCOT beziehen (siehe Kap 4.2). Dadurch wird auch deutlich, dass die Bedeutung und Verwendung des Rollenkonzepts in CTA sich an diesen Fachbereichen orientiert, um CTA-Positionen auszuarbeiten. Beim Rollenkonzept

der Actors handelt es sich um einen für den Fachbereich der Soziologie typischen Usus, der in CTA übernommen wird, um soziologischen Konventionen zu entsprechen und anschlussfähig an die soziologische Debattenkultur zu sein. Aufgrund der etablierten Verwendung in der Soziologie wird in CTA mit dem Rollenkonzept keine Notwendigkeit verbunden, über seine Verwendung weiterreichend zu reflektieren. Entsprechend wenig Aufmerksamkeit wird in der CTA-Literatur der aktiven Auseinandersetzung mit dem Rollenkonzept gewidmet.

Eine weitere Verbindung zwischen beiden Rollenkonzepten der Involvierten und somit zwischen den beiden Ansätzen CTA und VSD stellt die Actor-Network-Theory (ANT) dar (Akrich, 2023; Callon, 1984; Latour, 1987; Law, 1992). Eine zentrale Position der ANT besteht darin, dass soziale Prozesse und Phänomene nicht ausschließlich auf die Aktivität von Menschen zurückgeführt werden, sondern dass auch die Möglichkeit einer Beteiligung nicht-menschlicher Entitäten in Erwägung gezogen wird. Hieran wird eine direkte Verbindung zur zuvor erläuterten Bedeutung der Stakeholder in VSD ersichtlich, wo auf eben jene nicht-menschlichen Entitäten Bezug genommen wird. In der ANT ermöglicht es diese Position, Verbindungen zwischen semiotischen Elementen (wie z.B. Konzepten oder Ansätzen) sowie materiellen Elementen (wie z.B. technische Artefakte oder Gebäude) zu untersuchen. Darin liegt sowohl für VSD wie auch für CTA ein bedeutsamer Mehrwert, da für das in ihnen verfolgte Erkenntnisinteresse besonders Aspekte zu mehr abstrakten Konstrukten wie Wertefragen und Gesellschaftsdynamiken von zentraler Bedeutung sind. In der ANT kommt ferner eine semiotische Auslegung des Actor-Rollenkonzepts als „Actant“ zum Tragen, der eine Entität beschreibt „that engages in action or is granted the ability to do so by other entities“ (Rafael, 2023, S. 515). Auch diese Interpretation des Rollenkonzepts ist in der Literatur zu beiden Ansätzen wiederzuerkennen. Die ANT eröffnet für die Zwecke in CTA und VSD zudem weitere wertvolle Perspektiven auf Gestaltungsprozesse, da in der ANT soziale Prozesse und Phänomene als zusammenhängende und sich wechselseitig bedingende Netzwerkstrukturen betrachtet werden. Die Netzwerk-Analogie lässt sich auch auf Gestaltungsprozesse übertragen, beispielsweise um die Zusammenhangsstrukturen und Abhängigkeitsverhältnisse der als relevant identifizierten menschlichen und nicht-menschlichen Akteure bzw. Aktanten zu untersuchen (Green et al., 1999). Der Betrachtung von Gestaltungsprozessen als Netzwerkstrukturen liegt zudem eine grundsätzliche höhere Komplexität zugrunde (z.B. Abhängigkeitsstrukturen, Ursache-Wirkungs-Mechanismen) und erfüllt damit eine weitere in CTA und VSD zum Ausdruck gebrachte Anforderung. Sie bricht die mehr linearen und gerichteten Vorstellungen von Input-Output-Verhältnissen auf, wie sie in mehr auf die Technik fokussierten Ansätzen zum Tragen kommen. In Verbindung mit der bewusst angestrebten Flexibilität der Rollenkonzepte der Involvierten wird es beiden Ansätzen so möglich, die Komplexität von Gestaltungsprozessen für ihre Zwecke deutlich ganzheitlicher zu erfassen sowie die Ausschnitte sozialer Realität, die in CTA und VSD betrachtet werden, deutlich detaillierter abzubilden. Für CTA liegt hierin der besondere Mehrwert, Gestaltungsprozesse als andauernde, sich ständig weiterentwickelnde ko-evolutionäre soziale Prozesse zu konzipieren und untersuchen (Letch, 2010). Ein weiterer Hinweis der nahelegt, dass die Interpretation und Verwendung des Rollenkonzepts der Actors besonders in CTA Paralle-

len zu den Bedeutungen und Verwendungen in der ANT besitzt, liegt darin, dass Arie Rip etwa zeitgleich zum Beginn der Arbeiten zur Entwicklung von CTA in den 1980er-Jahren sich gemeinsam mit Michel Callon und John Law zu den „Dynamics of Science and Technology“ auseinandergesetzt hat (Callon et al., 1986).

### 5.5.2 Die informierten Einflussnehmer

Die „Rollenkonzeptpaare“ der informierten Einflussnehmer (z.B. Impactors und Impactees) dienen in CTA dazu, anhand von Gegenüberstellungen Missstände und Gegensätze in Gestaltungsprozessen aufzuzeigen und diese durch die in CTA vorgeschlagenen Methoden zu adressieren (Gewinner-Verlierer-Metapher). Die Untersuchung der Ursprünge der informierten Einflussnehmer zeigt, dass in den Fachbereichen der Managementlehre, sowie punktuell in den Fachbereichen der Soziologie und Nanotechnologie ähnliche Gedanken vorhanden und etabliert sind.

#### Impactors

Die Hintergründe zur Verwendung der Impactors in CTA sind primär auf linguistischer Ebene verortet. Ähnlich wie das „constructive“ weist das „impact“ eine Bedeutungspluralität auf, die in CTA verschiedentlich ausgeschöpft wird. Mit dem Verb „impact“ wird üblicherweise zum Ausdruck gebracht, dass Personen, Gegenstände oder weitere Entitäten auf eine definierte oder unbestimmte (zusätzliche) Entität eine Auswirkung ausüben. Diese Bedeutung wird in CTA verwendet, um einerseits die Auswirkungen technologischer Innovationen auf das gesellschaftliche Leben zu beschreiben (societal impacts), sowie um dem Selbstanspruch Ausdruck zu verleihen, durch die Anwendung von CTA-Methoden positive Auswirkungen in und auf Gestaltungsprozesse auszuüben. Im Zusammenhang mit dem Rollenkonzept der Impactors erfährt es die Bedeutung, die verschiedenen Positionen von an Gestaltungsprozessen beteiligten Akteuren zu beschreiben und die damit zusammenhängenden Auswirkungen (impacts) zu antizipieren – einem zentralen Ziel in CTA. Wesentlicher Fokus liegt dabei auf der Hervorhebung des Wissensvorsprungs über den Impactors typischerweise verfügen und damit zusammenhängend dem Machtungleichgewicht zu Impactees, dessen Ausgleich mit CTA-Methoden angestrebt wird.

Bei weiterer Auseinandersetzung mit der CTA-Literatur auf linguistischer Ebene findet sich darüber hinaus die Verwendung der Bezeichnung von „impactors“ in weiteren Sinnzusammenhängen wieder, insbesondere im Rahmen technischer Darlegungen die aus der Nanotechnologieforschung stammen. Hier werden mit Impactors jedoch keine Personen bezeichnet, sondern verfahrenstechnische Geräte zur Bestimmung von Partikelmassen in Gasen:

*„One method used to sample for nanoparticles is the low-pressure nano-cascade impactor, which uses five impactor plates from sizes between 10 to 100 nm“*  
(Renn & Roco, 2006, S. 18).

Bei einem Kaskadenimpaktor handelt es sich dabei um ein spezielles Partikelmessgerät, das häufig zur Erfassung von Staubpartikeln in Gasvolumen eingesetzt wird. Diese Bedeutung ist im Zusammenhang mit der Verwendung des Rollenkonzepts der Impactors in CTA auffällig, da neben der offensichtlichen übereinstimmenden Bezeichnung, die Nanotechnologie eben jenes Technologiefeld darstellt, in dem CTA ab etwa Mitte der 2000er-Jahre verstärkt angewendet und weiterentwickelt wurde. Etwa zeitgleich beginnt die Verwendung des Rollenkonzepts der Impactors in der CTA-Literatur einzusetzen und wird ab diesem Zeitpunkt in der konsekutiven CTA-Literatur verwendet. Dies lässt den Schluss zu, dass die Verwendung des Rollenkonzepts der Impactors vom in der Nanotechnologie üblichen Sprachgebrauch inspiriert wurde.

### Insider und Outsider, Enactors und comparative Selectors

Eine unmissverständlich zum Ausdruck gebrachte Übernahme von Rollenkonzepten in CTA stellen die beiden Rollenkonzeptpaare der Encators und comparative Selectors sowie der Insider und Outsider dar. Diese stammen aus einer Publikation der beiden Ökonomen Raghu Garud und David Ahlstrom (Garud & Ahlstrom, 1997). In dieser Arbeit setzen sich die beiden Autoren mit den verschiedenen Ansätzen auseinander, die Wissenschaftler verwenden, um Technologien zu bewerten (assess). Sie unterscheiden dabei grundlegend zwischen Insidern und Outsidern, wobei sie mit Insidern Wissenschaftler bezeichnen, die direkt an der Entwicklung von Technologien beteiligt sind, während Outsider Wissenschaftler bezeichnen, die Technologien primär evaluieren, fördern oder regulieren, ohne direkt in ihre Entwicklung involviert zu sein. Sowohl Insider wie auch Outsider besitzen nun verschiedene Methoden und Ansätze, die sie ausgehend von ihrer relativen Position zu einer Technologie anwenden, um diese zu beurteilen. Ausgehend von dieser grundlegenden Prämisse beschreiben Garud und Ahlstrom das Konzept der enactment cycles und selection cycles, die auch in CTA als zentrales gedankliches Referenzkonstrukt dienen. Bei den enactment cycles handelt es sich um Dynamiken, die durch die Aktivierung der Evaluationsansätze von Insidern aktiviert werden. Als Illustrationsbeispiel hierfür dienen Szenarien, die es Insidern ermöglichen, Zukunftsvisionen zu formulieren und mit ihrer Hilfe die notwendigen Schritte zu identifizieren, die das Erreichen dieser Zustände erfordern (Garud & Ahlstrom, 1997, S. 41). Die beiden Autoren fahren fort zu beschreiben, dass die Insider jedoch im Verlauf dieses Prozesses dazu nei-

gen, überwiegend die positiven Aspekte ihrer Zukunftsvisionen in den Vordergrund zu stellen und mögliche negative Effekte in Zukunftsszenarien auszublenken. Enactment cycles stellen im Denken der beiden Autoren selbstreferenzielle Prozesse dar, die ab einem bestimmten Zeitpunkt eine Eigendynamik entwickeln und in deren Verlauf Insider sich sukzessive zu Enactors wandeln, die ihre Funktion zunehmend in der Durchsetzung bestimmter Entscheidungen oder Entfernung bestimmter Hindernisse interpretieren, um die von ihnen entwickelten imaginären Zielzustände zu Realitäten umzuformen. Dieser Wandlungsprozess wird von den Autoren als fließend beschrieben, wodurch das Rollenkonzept der Encators trotz der dargelegten Unterschiede in der Arbeit von Garud und Ahlstrom wie auch in der CTA-Literatur häufig als Synonym verwendet wird.

Im Gegensatz zu Insidern verfolgen Outsider die Strategie von selection cycles und besitzen, verglichen mit Insidern, typischerweise einen weiteren Betrachtungshorizont:

*„The global scans outsiders engage in automatically lead them to adopt an epistemology that is consistent with large sample statistics“* (Garud & Ahlstrom, 1997, S. 41).

Sie betrachten im Gegensatz zu Insidern mehrere mögliche Zukunftsentwicklungen und verschiedene technologische Möglichkeiten zur Realisierung dieser, weshalb sie häufig ein gemeinsames Kriterium formulieren oder finden müssen, anhand dem sich eine betrachtete Technologie mit weiteren Technologien oder Anwendungen messen und vergleichen lässt. Die Bemühungen um eine vergleichende Bewertung von Technologien geschieht vor dem Hintergrund, eine Rangfolge von Technologien zu erarbeiten, deren Bewertung anhand von übertragbaren Kriterien erfolgt. Im weiteren Verlauf soll basierend auf derart herbeigeführten Bewertungen Kriterien (z.B. Listen) entwickelt werden, mit deren Hilfe die (Weiter-)Förderung oder Förderungsstopps bestimmter technischer Entwicklungen legitimiert werden können. Auch Outsider wandeln sich somit im Verlauf ihrer Arbeit hin zu comparative Selectors, womit auch zwischen diesen beiden Rollenkonzepten eine gewisse synonyme Bedeutung besteht, die so auch in CTA zum Tragen kommt. Den Zusammenhang von enactment cycles und selection cycles stellen die beiden Autoren auch visuell dar (siehe Abb. 42).

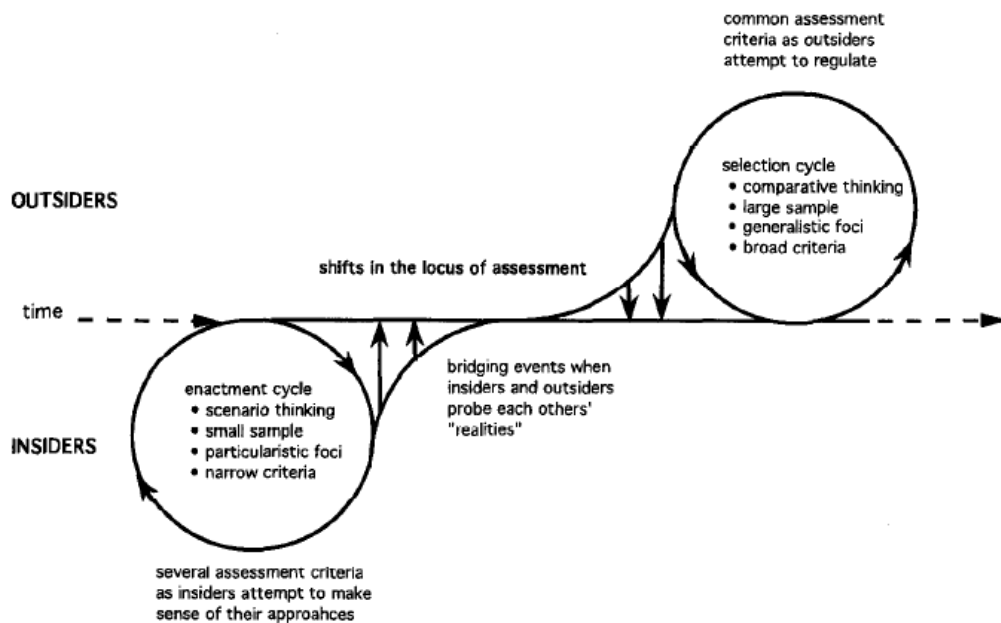


Abbildung 42: Zeitlicher Zusammenhang von enactment cycles und selection cycles nach Garud und Ahlstrom (1997, S. 44)

In Ergänzung zu ihren Ausführungen der Übernahme dieser Rollenkonzepte in CTA, weisen Rip und Robinson (2013, S. 43) zudem darauf hin, dass der Begriff des „enactment“ aus der soziologischen Theorie des symbolischen Interaktionismus stammt. So legen beispielsweise die beiden Soziologen Dennis und Martin (2005) in einer Arbeit dar, dass sich viele Untersuchungen des symbolischen Interaktionismus mit der Entstehung und Inkraftsetzung (enactment) von Machtstrukturen und ihrer weitreichenden Auswirkung auf Gruppen und Individuen befassen. Rip und Robinson legen dar, dass Garud und Ahlstrom diese Interpretation von enactment in ihrer Arbeit auf Entscheidungsträger in TG-Prozessen übertragen, um zum Ausdruck bringen zu können, welche Handlungen diese zur Realisierung ihrer Ziele und Visionen durchführen. Eine weitere Ergänzung zur Übernahme der Rollenkonzepte von Garud und Ahlstrom findet sich in einer Arbeit von Parandian (2012, S. 11), wonach es sich bei den von Insidern zu Enactors gewandelten Gruppen und Individuen um die primäre Zielgruppe der CTA-Methoden handelt. Besonders Ihnen soll durch die Anwendung von CTA-Methoden ein Reflexionsraum eröffnet werden, der ihnen die vielfältigeren technologischen Alternativen und Zukünfte gewahr werden lässt.

### 5.5.3 Die Unbeteiligten und Betroffenen

Auch die Rollenkonzepte der Unbeteiligten und Betroffenen zählen zu den primären Zielgruppen in CTA und VSD. Die stärkere Fokussierung und Auseinandersetzung mit den Rollenkonzepten der indirect Stakeholders und Impactees ist eine zentrale Motivation für die Entwicklung und Anwendung beider Ansätze. Diese Motivation fußt auf der in beiden Ansätzen zum Ausdruck gebrachten Kritik, dass den Unbeteiligten und Betroffenen in

Gestaltungsprozessen typischerweise zu wenig Aufmerksamkeit gewidmet wird, wenngleich nach in VSD und CTA vertretenen Ansichten zu verschiedenen Zeitpunkten und an verschiedenen Stellen in Gestaltungsprozessen entsprechendes Wissen und entsprechende Möglichkeiten zur Verfügung stünden. Diese Kritik richtet sich grundsätzlich an sämtliche an Gestaltungsprozessen beteiligten Disziplinen, wodurch es sich jedoch auch um eine diffuse Kritik handelt, da sie sich an eine mehr allgemeine Adressatengruppe richtet, deren Adressierung in tendenziell viele (interdisziplinäre) Zuständigkeitsbereiche fällt. Für CTA ist diese Kritik nicht ungewöhnlich, da es einen Ansatz darstellt, der in TA entwickelt wurde, wo häufig Kritik an TG-Disziplinen geübt wird (und Communitymitglieder in TA sich selbst *nicht* als einer solchen Disziplin zugehörig verstehen). Im Vergleich dazu handelt es sich bei VSD bei dieser Kritik mehr um eine „nach innen“ gerichtete, da sich Vertreter von VSD, im Unterschied zu Vertretern von CTA, selbst als Teil der TG-Disziplinen und somit als Adressaten der Kritik verstehen. Beide Ansätze streben danach, diese identifizierte Lücke zu schließen und das vorhandene Wissen über eine stärkere Auseinandersetzung mit indirect Stakeholdern und Impactees aufzugreifen und insbesondere die methodischen Möglichkeiten hierfür weiterzuentwickeln. In beiden Ansätzen ist die Schaffung von Partizipationsmöglichkeiten für Unbeteiligte und Betroffene ein primäres Ziel, zu dessen Realisierung sie jedoch unterschiedliche Zugänge und Perspektiven wählen. In VSD erfolgt hierfür ein Rückgriff auf Erkenntnisse der Nutzerforschung der Communities HCI und PD, während in CTA vermehrt technikhistorische Erkenntnisse und Erkenntnisse aus TA aufgegriffen und verwendet werden. Was die Charakterisierung und Verwendung der Rollenkonzepte der Unbeteiligten und Betroffenen in beiden Ansätzen eint, ist ihre gedankliche Betrachtung vom Ende bzw. Ergebnis eines Gestaltungsprozesses. Ferner beziehen sie sich sehr stark auf die entsprechenden Gegenstück-Rollenkonzepte der direct Stakeholders und Impactors, da viele der konkreten Eigenschaften und Positionen der indirect Stakeholders und Impactees häufig im starken Kontrast zu den Eigenschaften und Positionen ihrer Gegenstück-Rollenkonzepte stehen. In beiden Ansätzen wird zudem als eine ihrer zentralen Positionen in Anspruch genommen, dass Unbeteiligte und Betroffene in sämtlichen Gestaltungskontexten vorhanden sind. CTA und VSD zeichnen sich dadurch aus, Wege zu finden oder zu schaffen, den Positionen von Impactees und indirekten Stakeholdern mehr Aufmerksamkeit in Gestaltungsprozessen zu widmen und sie methodisch aktiv und systematisch in Gestaltungsprozesse einzubinden. Darin liegt eine der zentralen Innovationsfelder beider Ansätze sowie ein wichtiger Anknüpfungspunkt zur MKZG-Debatte in HCI.

## Indirect Stakeholders

Hintergrund der indirect Stakeholders ist die in VSD identifizierte Notwendigkeit der Erweiterung des klassischen Nutzerverständnisses in Gestaltungsprozessen:

*„a good deal is known and documented about direct stakeholder roles in the human-computer interaction literature. Less has been documented about indirect stakeholder roles“* (B. Friedman & Hendry, 2019, S. 41).

Da es sich beim Rollenkonzept der indirekten Stakeholder um eine Präzisierung bzw. um ein Derivat des Stakeholder-Rollenkonzepts handelt, gelten viele der bereits erläuterten Hintergründe zu Stakeholdern (siehe Kap. 1) auch für die indirect Stakeholder. Eine zentrale Referenz für die Ausarbeitung des Rollenkonzepts in der VSD-Literatur sind häufig die Stakeholder- und Nutzerverständnisse, die in den Fachbereichen der HCI und des PD zum Tragen kommen. Diese werden häufig dazu verwendet, um ihre Defizite in Bezug auf indirekt betroffene Gruppen und Individuen darzulegen und diese zu erweitern, besonders bezüglich Effekten der Gestaltung und der Nutzung von technischen Systemen. Ausgehend von dieser Kritik liegt die Bestrebung in VSD darin Möglichkeiten aufzuzeigen, wie die Methoden zur Ermittlung direkter Stakeholder auch dazu verwendet werden können, indirekte Stakeholder zu identifizieren, zu beschreiben und sie in einen Gestaltungsprozess miteinzubeziehen. Dabei wird herausgearbeitet, dass die Identifikation von indirekten Stakeholdern häufig mit nur einem geringen Mehraufwand verbunden ist, jedoch großen Mehrwert hinsichtlich der Mensch- und Werteorientierung von Gestaltungsprozessen bietet.

## Impactees

Bei den Impactees handelt es sich um ein für TA-Arbeit typisches Rollenkonzept. TA-Arbeit besteht häufig darin, die Chancen und Risiken von Technologien systematisch zu untersuchen, wodurch besonders in der Auseinandersetzung mit technologischen Risiken potenziell negativ betroffene Gruppen und Individuen identifiziert werden. So werden in TA beispielsweise die Problemdefinitionen und Lösungsansätze von „Entscheidern und Betroffenen“ (Bösch et al., 2018, S. 87) kontrastiert, oder die Folgen von Entscheidungen untersucht, die sich auf „beteiligte Akteure als auch potentiell Betroffene“ (Weyer, 1994, S. 38) auswirken. Die Unterscheidung von Entscheidungsträgern und Betroffenen ist dabei eine mehr soziologische Betrachtung von Mensch-Technik-Verhältnissen, die in TA von Communitymitgliedern mit sozialwissenschaftlicher Expertise eingebracht werden. Somit gehört das Rollenkonzept zum „Standardrepertoire“ von TA-Arbeit, jedoch wird es neben TA auch in benachbarten Feldern, wie beispielsweise der Innovationsforschung verwendet. So beschreiben Rubenstein und Geisler (1988) in einer Arbeit, dass sich die Identifikation von Impactees typischerweise als Herausforderung darstellt, da sie neben ihrer Rolle als Impactees auch die Rolle von Kunden ausüben. Ergänzend beschreiben die beiden Autoren, dass Impactees typischerweise an verschiedenen Stellen des politischen und sozialen Spektrums verortet sind, was ihre Identifikation zusätzlich erschwert.

Ein weiterer Grund für die Verwendung von Impactees in CTA ist mit den historischen Hintergründen des Ansatzes als solchem verbunden. Johan Schot (2001) erläutert an einer Stelle in der Literatur die Entwicklungsgeschichte von CTA, die mit der Erkenntnis beginnt, dass ein erfolgreiches Technologiemanagement in den Niederlanden zusätzlich zu TA und Gestalten die Beteiligung von betroffenen Akteuren aus der Gesellschaft erforderte („broadening“):



*„The initiators of CTA based their ideas on analysis of historical cases of broadening design (or lack of broadening) and theories of technical change drawn from evolutionary economics and the history and sociology of technology“* (Schot, 2001, S. 40–41).

Die daraus resultierende Perspektive auf TG-Prozesse verbindet technikhistorische, soziologische und ökonomische Erkenntnisse zu einem evolutionären Modell (co-evolution of science, technology and society), in dem mit den Impactees besonders die betroffenen Gruppen und Individuen aus der Gesellschaft frühzeitig mit Entscheidungsträgern aus Politik und Wirtschaft zusammengeführt werden sollen. Insofern repräsentiert das Hervorheben und Betonen des Rollenkonzepts der Impactees in besonderer Weise die Bemühungen in CTA, TG-Prozesse und ihre soziale Einbettung konzeptionell möglichst ganzheitlich zu fassen.

#### 5.5.4 Die Gestalter

Die Rollenkonzepte der Gestalter beziehen sich auf gängige Beschreibungen von Aufgaben- und Berufsprofilen der Gruppen und Individuen, die im gesellschaftlichen Leben technische Systeme unter Anwendung ihrer kognitiven und handwerklichen Fähigkeiten realisieren. Sich mit den Aufgaben und Fähigkeiten von Entwicklern, Ingenieuren und Designern auseinanderzusetzen stellt eine Notwendigkeit in beiden Ansätzen dar, die in VSD im Vergleich mit CTA eine etwas höhere Priorität besitzt, da sich der Ansatz deutlich expliziter an diese Zielgruppen richtet und sie in der Ausführung ihrer Tätigkeiten zu unterstützen sucht. Die Verwendung der Rollenkonzepte der Gestalter stellt dabei in sämtlichen Disziplinen, die sich mit TG und der Untersuchung von Mensch-Technik-Verhältnissen auseinandersetzen eine gängige Praxis dar, einen Konsens der sich zur Vermittlung und zum Austausch von Wissen in der Vergangenheit etabliert hat. Somit handelt es sich bei Verwendung der Rollenkonzepte in CTA und VSD um die Übernahme einer gängigen Konvention, um anschlussfähig an die für CTA und VSD relevanten Debatte-kulturen zu sein. Die Besonderheit der Verwendung der Rollenkonzepte erschließt sich durch die Ausprägungen und Attribute, die die Rollenkonzepte in CTA und VSD erfahren, da sich die Rollenkonzepte dort im Vergleich mit anderen Gestaltungsansätzen deutlich intensiver mit Fragen zu der mit gestalterischem Handeln verbundenen Verantwortung auseinandersetzen. In beiden Ansätzen ist eine konkrete Unterscheidung von Entwicklern, Ingenieuren oder Designern häufig jedoch wenig ausschlaggebend, da sie in weiten Teilen der Literatur schlicht stellvertretend für technische Expertise in einem allgemeinen Sinne stehen, deren konkrete Ausprägungen sich durch die Eigenschaften eines Gestaltungskontexts und den dort aktiven und betroffenen Gruppen und Individuen ableiten („Kontextsensitivität“).

Dass in VSD im Vergleich zu CTA häufiger von Designern die Rede ist, geht auf eine Tradition der HCI-Community zurück. In der HCI-Community verstehen sich viele Communitymitglieder als schaffende Professionals in konzeptuellen wie auch in künstlerischen Bereichen (z.B. Interaktionsdesign, Interfacedesign), und haben häufig eine entsprechende Berufsausbildung in traditionellen Gestaltungsberufen wie der Mediengestaltung oder

dem Produktdesign absolviert. Aufgrund verschiedener Ausbildungsschwerpunkte zu gestalterischen Tätigkeiten auf technischer Ebene (z.B. Programmierung, Konstruktion) in solchen Berufsprofilen hegen manche der Communitymitglieder den Wunsch, ihre Kompetenzen transparent zu kommunizieren, um möglichen Missverständnissen zu Erwartungshaltungen ihrer Fähigkeiten vorzubeugen. Häufig sind die Gruppen und Individuen, die in HCI und entsprechend in VSD als Designer bezeichnet werden, durchaus auch technisch versiert, jedoch liegt der Schwerpunkt ihrer Aufgaben und Kompetenzen typischerweise deutlich stärker in der konzeptuellen Gestaltung und technischen Realisierung von beispielsweise Möglichkeiten zur Navigation in einem System (im Fachjargon häufig mit „Frontend“ bezeichnet), wie der technischen Realisierung von beispielsweise der dafür notwendigen Dateninfrastrukturen („Backend“). In der HCI wird deshalb häufig mit Werkzeugen zur Visualisierung von Gedanken, Ideen und Systemkonzepten gearbeitet, wodurch viele Communitymitglieder mit Erkenntnissen wie beispielsweise den Gestaltungsgesetzen der Wahrnehmung (z.B. Bühler et al., 2017) und weiteren Erkenntnissen der Ästhetik und visuellen Kommunikation vertraut sind. Dadurch bestehen in den Aufgaben- und Kompetenzprofilen von Communitymitgliedern in HCI durchaus Überschneidungen zu jenen von beispielsweise Grafikern, Architekten, Modedesignern oder Künstlern. Für die HCI-Community wie auch in VSD ist mit dem Rollenkonzept der Designer daher die Funktion einer potenziellen Reichweitenerhöhung verbunden, denn durch die Ansprache sämtlicher gestaltender Berufe werden potenziell sämtliche Gruppen und Individuen angesprochen, die sich sämtlichen dieser Berufsgruppen zugehörig betrachten.

### 5.5.5 Die intervenierenden Prozessbegleiter

Die CTA agents stellen eine Anlehnung an das Rollenkonzept der Change Agents aus der Organisationsforschung dar, welche in der CTA-Literatur auch offen kommuniziert wird (Rip & Robinson, 2013, S. 47). In der Organisationsforschung werden Change Agents auf verschiedene Weise charakterisiert, grundsätzlich handelt es sich dabei jedoch um Gruppen oder Individuen, die in Organisationen Veränderungen herbeiführen und dazu Auswirkung auf sämtliche Prozesse einer Organisation nehmen können (Lunenburg, 2010). Häufig handelt es sich dabei um eines von mehreren Aufgabenfeldern von Personen mit Führungsaufgaben, jedoch können auch Personen die keine Führungsposition bekleiden mit den Aufgaben von Change Agents betraut sein. Häufig agieren Change Agents auch als (wissenschaftliche) Berater von Organisationen in Strukturfragen, was, wie auch weitere Eigenschaften von Change Agents, Parallelen zum traditionellen Selbstverständnis vieler TA-Arbeit verrichtenden Gruppen und Individuen aufweist. Typischerweise stehen Change Agents in Kontakt mit Verantwortungsträgern einer Organisation, tragen kritische Informationen zu Entscheidungsprozessen bei und beteiligen sich an der Auflösung von Konfliktsituationen. Sie identifizieren Chancen und Risiken in den Abläufen von Organisationen und sind häufig verantwortlich für die Weiterentwicklung eines Prozesses oder Produktes in einer Organisation. Weitere typische Charakteristika die mit Change Agents verbunden werden sind eine ausgeprägte Kommunikationsfähigkeit, Empathie für ihre Mitmenschen sowie ein hohes Maß an Resilienz und Durchsetzungsvermögen, das sie zur Realisierung der von ihnen angestrebten Veränderungen häufig benö-

tigen. Die Geschichte des Rollenkonzepts innerhalb der Organisationsforschung zeichnet sich durch viel Bedeutungswandel aus, und wird auf den Sozialpsychologen Kurt Lewin zurückgeführt (Ottaway, 1983). Lewins Forschungsinteresse galt Veränderungsprozessen innerhalb von Gruppen in Organisationen, speziell der Frage, wie solche Gruppen Konflikte lösen (Lewin, 1947). Seit der erstmaligen Formulierung von Change Agents Ende der 1940er-Jahre, haben die mit dem Rollenkonzept verbundenen Bedeutungen viele Veränderungen und Hypezyklen durchlaufen, wie etwa während der 1990er-Jahre, als der Begriff im Kontext von Unternehmensberatungen und Consultingfirmen und entsprechend auch in der Organisationsforschung erneut erhöhte Aufmerksamkeit erfuhr (z.B. Fougere, 1991; Markus & Benjamin, 1996; Simonsen, 1997). Dies ist auch die Zeit, in der CTA erste fortgeschrittene Entwicklungsstadien erreicht. In Verbindung mit der Tatsache, dass in der Ausarbeitung von CTA auch die Erkenntnisse der Organisationslehre Beachtung erfahren (siehe Kap. 5.5.2), ist es naheliegend, dass auch das damals viel diskutierte Rollenkonzept der Change Agents entsprechende Beachtung erfährt.

Durch ihr Aufgabenprofil sind Change Agents häufig auch als Beobachter und Begleiter von Projekten und Prozessen tätig. Darin liegt ihre zentrale Funktion bzw. Interpretation der CTA agents in CTA. Wie Rip und Robinson (2013) beschreiben, sollen CTA agents Veränderungsprozesse mit der Zielsetzung begleiten, eine stärker reflektierte und inklusive TG herbeizuführen. Ihre Hauptaufgabe besteht in der Durchführung von punktuellen Interventionsmaßnahmen in Gestaltungsprozessen. Ähnlich eines wissenschaftlichen Gremiums das – wie in vielen TA-Kontexten – mit Begleitforschung beauftragt ist, sollen CTA agents dazu die in CTA vorgeschlagenen Methoden anwenden und Reflexions- und Diskussionsräume eröffnen. Auch sollen sie politische und wirtschaftliche Entscheidungsträger teilweise direkt, teilweise indirekt bei ihrer Entscheidungsfindung beraten, jedoch verfügen sie im Verständnis in CTA selbst nur über geringe Haftungsansprüche für die Resultate oder (Miss-)Erfolge von Gestaltungsprozessen. Ihre Verantwortung beschränkt sich auf die Zusammenführung von Akteuren aus Politik, Wirtschaft und Gesellschaft, sowie der zielführenden Durchführung der Interventionsmaßnahmen. Somit entspricht die Vorstellung von CTA agents in weiten Teilen dem traditionellen Selbstbild vieler TA-Forscher, die in der Politikberatung tätig sind.

### 5.5.6 Die Nutzer

Ähnlich wie bei den Rollenkonzepten der Gestalter, handelt es sich auch bei den Nutzern um ein Rollenkonzept, das einer Konvention in vielen Disziplinen entspricht, die sich mit TG auseinandersetzen. Nutzer stellen das Gegenstück zu Gestaltern als Anwender von technischen Systemen dar, die in den Überlegungen in Gestaltungsprozessen insofern von Relevanz sind, da sie die gestaltete Technik für bestimmte Zwecke verwenden können sollen. Sie sind die „Empfänger“ von Technik (Spiekermann, 2010). In vielen technischen und ökonomischen Fachbereichen, besonders auch in Kontexten der Gestaltung von IT-Systemen, ist mit dem Rollenkonzept der Nutzer die Vorstellung von Endverbrauchern (end user) verbunden. In diesem Kontext werden Gestaltungsprozesse als eine Wertschöpfungskette betrachtet, wobei Nutzer gedanklich häufig gleichgestellt werden mit den Konsumenten eines Produkts oder Personen, die eine Dienstleistung in Anspruch

nehmen. In vielen Gestaltungskontexten stellen Nutzer daher das Ende der Wertschöpfungskette dar, da sie ein technisches System käuflich erwerben um damit im beruflichen oder privaten Kontext spezifische Aufgaben durchzuführen. Stellenweise können Nutzer auch die Funktion von Auftraggebern bzw. Initiatoren von Gestaltungsprozessen erfüllen. Für sämtliche gestaltenden Disziplinen, ihren Gestaltungsansätzen und somit auch in CTA und VSD sind Nutzer deshalb aus einer Vielzahl von Gründen relevant. Die Interessen, Herausforderungen und Bedürfnisse von Nutzern stellen häufig die zentralen Bezugspunkte zu in der Realität existierenden Zuständen dar, die ein technisches System adressieren soll. Das Erheben und Nachvollziehen dieser Eigenschaften ist von hohem Stellenwert in Gestaltungsprozessen, da sich die Qualität, die Angemessenheit, oder auch die Effektivität eines technischen Systems dadurch auszeichnet, wie adäquat es diese Eigenschaften unterstützt. Aus den Charakteristika von Nutzern lassen sich funktionale, ergonomische, psychologische, soziale und weitere Gestaltungsanforderungen ableiten, die es in der Gestaltung gilt aufzugreifen und Möglichkeiten zu finden, sie angemessen zu adressieren. Neben den inhaltlichen und Zweckaspekten hängt von solchen Qualitätskriterien nicht zuletzt häufig auch der Verkaufserfolg eines Systems ab, einem weiteren sehr wichtigen Faktor in vielen Gestaltungsprozessen. Die Abgrenzung von Nutzern zu Gestaltern hat deshalb häufig pragmatische Hintergründe und soll es erleichtern, die hohe Komplexität von Gestaltungsprozessen (siehe Kap. 5.1) zu reduzieren und in der Gestaltungspraxis eine Arbeitsaufteilung zu ermöglichen. Diese Konvention wird in CTA und VSD übernommen, was es in den beiden Ansätzen ermöglicht, ihre Positionen auszuarbeiten oder beispielsweise auch Methoden zu Teilprozessen eines Gestaltungsprozesses zu entwickeln.

### 5.5.7 Die Technikenthusiasten und die Technikgegner

In den Rollenkonzepten der Proponents und Promoters kommen ähnliche Gedanken und Perspektiven zum Ausdruck, die aus der Managementlehre stammen und die die Entwicklung und Interpretation der Rollenkonzepte der informierten Einflussnehmer in CTA beeinflussen. Besonders das Rollenkonzept der Proponents stellt eine starke Anlehnung zu den von Insidern zu Enactors gewandelten Gruppen und Individuen dar (siehe Kap 5.5.2), die sich dadurch auszeichnen, dass sie sich für die Gestaltung bestimmter technischer Anwendungen einsetzen. Bei den Proponents, stärker noch bei den Promoters steht dabei im Vergleich mit anderen Rollenkonzepten in CTA mit der (wenig hinterfragten) Begeisterung für Technik eine ausgeprägte Charaktereigenschaft von Menschen im Vordergrund. Durch die Fokussierung auf die Überzeugung von Technik wird bei den Rollenkonzepten der Technikenthusiasten eine sehr geringe Abstraktionsebene angewendet, wodurch sie den Eindruck vermitteln, sehr spezifische Gruppen und Individuen zu beschreiben.

#### Proponents und Opponents

Wie bereits zuvor gewährt auch die Betrachtung der linguistischen Ebene der Technikenthusiasten aufschlussreiche Einblicke, um die Ursprünge der Rollenkonzepte zu ergründen. Im englischen Sprachgebrauch bezeichnet ein proponent eine Person, die sich öffentlich

für eine Sache, eine Idee, oder eine Handlungsoption (course of action) ausspricht und für diese argumentiert. Synonyme für proponents stellen advocats dar, das Antonym zu proponents sind die opponents, die in CTA ebenfalls als Rollenkonzepte verwendet werden. Im Vergleich zu beispielsweise den Rollenkonzepten der Gestalter oder der informierten Einflussnehmer handelt es sich bei diesem Rollenkonzeptpaar jedoch nicht um die Übernahme einer sprachlichen Konvention technischer oder ökonomischer Fachbereiche, sondern vielmehr um eine Konvention eines philosophischen Fachbereichs. Die Schule der dialogischen Logik wird auf den Philosophen und Mathematiker Paul Lorenzen zurückgeführt, der zusammen mit Kuno Lorenz die Dialogführung unter Menschen als ein Wechselspiel von verbalen Angriffen durch Proponents und ebenso verbalen Verteidigungen durch Opponents konzipierte (Lorenzen & Lorenz, 1978). Zu dieser Vorstellung von zwischenmenschlicher Debattenkultur hat Lorenzen seit den 1950er-Jahren verschiedene philosophische Essays verfasst, in denen er sich auch mit der Philosophie und Wissenschaftstheorie des Konstruktivismus auseinandergesetzt hat (Lorenzen, 1968, 1974). Er selbst gilt zudem als Begründer der Erlanger Schule, der eine methodologische Ausprägung des Konstruktivismus vertritt und versucht,

*„einen begründeten und zirkelfreien Aufbau der Wissenschaftssprachen und damit der Wissenschaften selbst zu leisten“ (Zitterbarth, 1991, S. 7).*

Durch die Auseinandersetzung mit konstruktivistischen Gedanken besteht somit zwischen der Arbeit von Lorenzen und CTA auch eine inhaltliche Verbindung, da sich CTA ebenfalls auf konstruktivistische Positionen gründet (siehe Kap 4.2). Die gesammelten Gedanken von Lorenzen werden gegen Ende der 1980er-Jahre unter dem Titel „Constructive Philosophy“ auch der englischsprachigen Philosophie zugänglich gemacht (Lorenzen, 1987), die in der CTA-Literatur nicht direkt referenziert wird, jedoch kann ihre inhaltliche, besonders ihre wissenschaftstheoretische Relevanz (philosophy of science) für CTA in der erweiterten TA-Literatur nachvollzogen werden. So wird beispielsweise in einer Arbeit von Decker (2001) beschrieben, dass im Kontext interdisziplinärer Debatten, wie sie in den CTA-Workshops angestrebt werden, wichtig sei, dass sämtliche Argumente allen Teilnehmern nachvollziehbar dargestellt werden müssen. Auch in einer Arbeit von Grunwald (2012) wird auf Lorenzen im Kontext der Auseinandersetzung mit Zukünften als soziale Konstrukte verwiesen. Dort wird hervorgehoben, dass die argumentative Stärke von Zukunftsvisionen nicht mit ihrer Eintrittswahrscheinlichkeit gleichzusetzen sei.

## Promoters und Opponents

Beim Rollenkonzept der Promoters kommt auch die linguistische Bedeutung des argumentativen Werbens für bestimmte Positionen zum Ausdruck, die sich in CTA speziell auf technische Anwendungen bezieht. Anders als bei den Proponents besteht bei den Promoters jedoch ein Bezug zu dem theoretischen Konstrukt des two-track-approaches in CTA, bei dem die Durchsetzung und gesellschaftliche Einbettung technologischer Innovationen den beiden Seiten der technology promotion und technology control (stellenweise als Synonym zu den comparative Selectors) zugeschrieben wird. Während sich die Institutionen, die der technology control zugerechnet werden, dabei primär auf legislative Einrichtungen beziehen, werden die Technik gestaltenden Institutionen als Promoters dargestellt:

*„This dichotomy between promotion and control of new technology is thus part of the de facto constitution of modern industrialised societies, and is reflected not only in the division of labour between government agencies, but also in cultural and political views, as in the assumption that there will be proponents and opponents to a new technology“* (Rip, 2008, S. 147).

Diese Vorstellung des two-track-approachs wird in der CTA-Literatur dargestellt als das Resultat einer Gesellschaftsform, die dem Dilemma begegnen sollte, Technologien als strategische Ressource im politischen und ökonomischen Wettbewerb einzusetzen und zu fördern, während andererseits Wege gefunden werden mussten, den Begleiterscheinungen und Folgen von Technikentwicklung und Technikeinsatz (z.B. Umweltverschmutzung, Arbeitsplatzsicherheit oder Schutz der Privatsphäre) zu begegnen (Rip et al., 1995). Dabei handelt es sich demnach auch um eine mehr historische Betrachtung der Zusammenhänge von Mensch und Technik, die zum Ergebnis hatte, dass sich in vielen modernen Gesellschaften in den 1980er-Jahren recht starre Top-Down-Strukturen entwickelt hatten, die die Entscheidungsgewalt über das Technologiemanagement besaßen. In CTA dient die Betrachtung dieser Entwicklung dazu, die mit den starren Strukturen verbundenen Schwachstellen aufzuzeigen und als Resultat das auch demokratisch motivierte Argument für mehr Beteiligung gesellschaftlicher Akteure an Entscheidungsprozessen zu motivieren. Die Schwachstellen des two-track-approachs haben sich nach Ansicht in CTA in der Geschichte bereits mehrfach offenbart, was bei Schot und Rip (1997) am Beispiel der Ludditen dargelegt wird. Die Ideale, die mit gesellschaftlichen Top-Down-Strukturen verbunden werden, sind nach Ansicht in CTA für zukunftsfähige Gesellschaftsmodelle keine Option, da es sich bei den Aneignungsprozessen von Technik um keine linearen Vorgänge, sondern um Vorgänge in Netzwerkstrukturen handelt, an denen viele Akteure an unterschiedlichen Stellen im gesellschaftlichen Leben beteiligt sind.

Als Argument für diese Position werden auch die vermehrt beobachtbaren gesellschaftlichen Widerstände zu technischen Entwicklungen angeführt, die in CTA typischerweise durch die Technikgegner als „Opponents of technology“ repräsentiert werden. Als wahrgenommene oder tatsächliche Verlierer solcher Entwicklungen ist der Ursprung dieses Rollenkonzepts in CTA stark verknüpft mit denen seiner Gegenstück-Rollenkonzepte der Proponents und Promoters. Die Opponents stellen die Gegenseite der verbalen Aushandlung zu potenziellen Nutzen von Technik dar und vertreten in dieser Debatte die Seite, die die Contra-Argumente zum Ausdruck bringt.

## 5.6 Blinde Flecke der Rollenkonzepte

Die Eigenschaften, Funktionen und Ursprünge der Rollenkonzepte zeigen, welche Facetten menschlichen Lebens und Handelns in CTA und VSD im Fokus stehen. Sie repräsentieren diejenigen Gruppen und Individuen, die in CTA und VSD Relevanz für die Gestaltung technischer Systeme besitzen und wie sie in den beiden Ansätzen in Bereiche sozialen Geschehens eingeordnet werden. Durch die mit ihnen verbundenen Interpretationen im Kontext spezifischer Ausschnitte sozialen Geschehens stellen sie zudem Grenzmarkierungen der Geltungsbereiche dar, die in beiden Ansätzen als gültig in Anspruch genommen werden. Sie fungieren als teilweise bewusst gesetzte, teilweise unbewusst gesetzte Abgrenzungen zu Diskursen in angrenzenden Fachbereichen, die menschliches Verhalten oder Gestaltungsprozesse aus anderer Perspektive betrachten. Durch die verschiedenen mit ihnen verbundenen Intentionen und Zwecke werden mit den Rollenkonzepten in CTA und VSD spezifische Aspekte menschlichen Lebens und Handelns fokussiert, während andere Gesichtspunkte nur am Rande betrachtet werden oder gänzlich außen vor gelassen werden. Dies hat auch der Blick auf die Ursprünge der Rollenkonzepte gezeigt, wonach sich manche Rollenkonzepte in CTA an Konventionen soziologischer Communities orientieren, während für viele Rollenkonzepte in VSD die Interaktionsbeziehung von Mensch und Technik deutlicher im Vordergrund stehen.

Die blinden Flecke der Rollenkonzepte, die Aspekte, die in und mit ihnen nicht betrachtet und behandelt werden, halten jedoch für die Beantwortung der übergeordneten Forschungsfrage der Dissertation wertvolle Hinweise bereit. Die Untersuchung der mit den gewählten Perspektiven auf menschliches Handeln verbundenen Betrachtungslimitierungen und Gedankengangslücken ermöglicht mir das Kontrastieren des bislang erarbeiteten Verständnisses der Rollenkonzepte in CTA und VSD und eröffnet dadurch auch Möglichkeiten und Potenziale, die zur Weiterentwicklung beider Ansätze verwendet werden können.

Durch die Untersuchung der blinden Flecke der Rollenkonzepte wird es mir möglich, das Verständnis der Rollenkonzepte zu vervollständigen und dadurch mögliche Anknüpfungspunkte, aber auch Defizite zwischen den beiden Ansätzen und dem MZKG-Diskurs in HCAI freizulegen. Meine Identifikation der blinden Flecke erfolgt dabei unter Berücksichtigung und Anerkennung der Geschichte beider Ansätze, sowie der mit den Rollenkonzepten verbundenen Zwecke und Ursprünge innerhalb der Gedankenwelten in CTA und VSD. Ich nehme jedoch vor dem Hintergrund der Fragestellung der Dissertation eine bewusst gewählte fragende Haltung ein, bei der ich die in den Ansätzen gesetzten Gültigkeitsgrenzen bewusst überschreite und sie mit relevanten Aspekten des MZKG-Diskurses konfrontiere.

### 5.6.1 Abstraktion vs. Spezifität

Beide Ansätze arbeiten vermehrt mit stark abstrakten, unspezifischen Rollenkonzepten wie beispielsweise den Involvierten oder den Unbeteiligten und Betroffenen. Die Gründe der Verwendung dieser Rollenkonzepte erschließt sich aus den Betrachtungsmaßstäben, die in beiden Ansätzen zur Betrachtung von Gestaltungsprozessen angewandt werden. So erfordert die Auseinandersetzung mit sehr frühen Gestaltungsphasen und technologischen Reifegraden in CTA die Arbeit mit gedanklichen Platzhaltern, um menschliches Handeln zum Ausdruck zu bringen, da zu Beginn der angestellten Betrachtungen und Untersuchungen häufig nur wenig Wissen über die tatsächlich involvierten und handelnden Gruppen und Individuen vorhanden ist. Es stehen häufig komplexe und vielschichtige Prozesse und Abhängigkeiten im Mittelpunkt, die eine hohe Intransparenz verbindet und die oftmals zuvor noch an keiner anderen Stelle versucht wurde zu überwinden. So geht aus der Literatur häufig auch hervor, dass auch nach der Identifikation konkreterer gesellschaftlicher Institutionen und Akteure aufgrund der hohen Dynamiken auf technischer und sozialer Ebene eine Opazität bestehen bleibt, wodurch sich die fortgesetzte Verwendung von Platzhalterrollenkonzepten in CTA-Arbeiten erklären lässt. Hinzu kommt, dass CTA- und VSD-Projekte häufig Prozesse begleiten, die lange Zeiträume umspannen und in deren Verlauf sich die für CTA und VSD relevanten Gruppen erst sukzessive und zeitlich verzögert formieren. Dadurch wird eine konstante Überprüfung und Reevaluation der als relevant identifizierten Gruppen und Individuen notwendig, um die angestrebten Gestaltungs- und Interventionsziele erreichen zu können. Aus diesem Grund wird in beiden Ansätzen wiederholt die Notwendigkeit für eine konstante Aufrechterhaltung der Aufmerksamkeit für die Dynamiken und Veränderungen im Feld eingefordert. Zusätzlich versuchen die Ansätze auf methodischer Ebene diesem Umstand gerecht zu werden, indem sie beispielsweise die Möglichkeit der mehrfachen Durchführung von Workshops oder Befragungen aufzeigen. Ein weiterer Grund für die Verwendung mehr abstrakter Rollenkonzepte geht darauf zurück, dass sowohl in VSD wie auch in CTA der Anspruch erhoben wird, auf die Gestaltung sämtlicher Technologien und somit auf sämtliche Gestaltungskontexte anwendbar zu sein. Um diesem Anspruch gerecht werden zu können, benötigen die verwendeten Rollenkonzepte eine hohe Flexibilität und Offenheit, weshalb sie wiederum einen gewissen Abstraktionsgrad und die damit verbundene Vagheit nicht unterschreiten können.



Hierin liegt jedoch ein potenzielles Risiko hinsichtlich der Nachvollziehbarkeit und auch der Umsetzbarkeit der vorgeschlagenen Methoden in CTA und VSD. Für potenziell interessierte Anwender, die zusätzlich zu ihren Erkenntnisinteressen mit einem konkreten gestalterischen Auftrag in Kontakt mit den Rollenkonzepten kommen, kann das Nicht-Unterschreiten von Abstraktionsgraden zu Verwirrung, Fehlinterpretation und Frustration führen, da sie keinen Bezug zu der ihnen vorliegenden, konkreten Gestaltungssituation herstellen können. Sie werden aufgrund der Vagheit der Rollenkonzepte indirekt mit der Aufgabe der Verrichtung von Transferarbeit auf ihr Gestaltungssetting beauftragt, die im Zweifel aufgrund fehlender Orientierungshilfen nicht erfolgen kann. Die VSD-Literatur wirkt diesem Risiko punktuell entgegen, indem dort mehrere Fallstudien referenziert werden, aus denen ersichtlich wird, dass beispielsweise obdachlose junge Menschen die Konkretisierung des Rollenkonzepts der direkten Stakeholder darstellen können, während staatliche Einrichtungen zur Unterstützung obdachloser junger Menschen die Konkretisierung des Rollenkonzepts der indirekten Stakeholder darstellen können. Die dennoch überwiegend geringe Verwendung von Fallstudien und Anschauungsbeispielen zur Illustration und Klarstellung der Rollenkonzepte in der Literatur zu beiden Ansätzen erschwert den Transfer und die Anwendbarkeit der Methoden, was angesichts des ansonsten hohen demonstrierten Grades an Selbstreflexion in beiden Ansätzen überrascht. In beiden Ansätzen kommt ein hohes Interesse bezüglich der Angemessenheit ihrer Rollenkonzepte zum Ausdruck, die jedoch durch fehlende Maßnahmen zur Prävention von Fehlinterpretationen an einer entscheidenden Stelle einen blinden Fleck aufweist. Dabei würde ein stärkeres Arbeiten mit Anschauungsmaterial zu Rollenkonzepten die Intentionen beider Ansätze unterstützen, um interessierten Anwendern eine inhaltliche sowie handwerkliche Hilfestellung anzubieten. Durch das explizite Ausweisen als Beispielausprägungen von Rollenkonzepten können bei interessierten Anwendern Lernprozesse unterstützt werden und auch die mit den Rollenkonzepten angestrebte Flexibilität bleibt erhalten.

### 5.6.2 Die Rolle des Individuums

In beiden Ansätzen wird menschliches Leben und Handeln aus unterschiedlichen Perspektiven und in verschiedenen Konstellationen betrachtet, jedoch wird der Mensch dabei nur selten als Individuum betrachtet. Während dieser Umstand in CTA etwas nachvollziehbarer ist, da der Ansatz im Vergleich zu VSD eine grundsätzlich höhere Abstraktionsebene zu etablieren und zu wahren sucht (Doorn et al., 2013, S. 237) und mehr mit der Vorstellung von Netzwerkstrukturen arbeitet, stellt er besonders in VSD einen blinden Fleck zu einigen der aufwandsreich aufgebauten Positionen dar. So wird in beiden Ansätzen anerkannt, dass jeder Gestaltungsprozess individuelle Charakteristika aufweist, der sie von anderen Gestaltungsprozessen unterscheidet. Dabei kann es sich, wie in der VSD-Literatur zur Sprache kommt, beispielsweise um den Umstand handeln, dass in zwei Gestaltungsprozessen derselbe Wert (z.B. Autonomie) im Vordergrund steht, dass jedoch die mit dem Wert verbundenen Interpretationen in Prozess A sich signifikant von jenen in Prozess B unterscheiden. Auch weiß besonders die VSD-Literatur um die damit verbundenen Hintergründe, denn Individuen bewerten Fragen und Aspekte in Gestaltungs-

prozessen unterschiedlich und greifen bei Prioritäts- und Wertefragen auf ihre persönlichen Erlebnisse und daraus entwickelten Maßstäbe und Argumentationsstrategien zurück. Diese individuellen Positionen und Strategien sind häufig eine Ursache von Wertekonflikten, die besonders in VSD von zentraler Bedeutung sind, und deren Analyse und Auflösung folglich die Hinwendung zur Betrachtung von Menschen als Individuen erfordern kann. Auch die unterschiedlichen Machtgefüge, die in einigen der Rollenkonzepte in CTA zur Sprache gebracht werden, erfordern letztlich häufig die Auseinandersetzung mit Individuen und individuellen Positionen, da diese über sehr viel oder aber sehr wenig Entscheidungsmacht verfügen. Somit könnten auch die Anstrengungen in CTA von der punktuellen und expliziten Betrachtung des Menschen als Individuum ein bislang ungenutztes Potenzial heben. In beiden Ansätzen besteht ein wesentlicher Argumentationsbestandteil darin, genau diese Unterschiede und Besonderheiten in Erfahrung zu bringen, und ihnen während des gesamten Gestaltungsprozesses hohe Aufmerksamkeit zu widmen, jedoch ohne dafür das Rollenkonzept des Individuums zu verwenden. So ist beispielsweise in der CTA-Literatur zu lesen, dass in CTA-Workshops die von Individuen zum Ausdruck gebrachten Ansichten sich häufig als Schlüsselaspekte für die Weiterentwicklung einer Technologie erweisen. Das Zugeständnis und die Offenheit, menschliches Leben und Handeln aus der Perspektive des Individuums zu betrachten, unterstützt die Realisierung der in CTA und VSD formulierten Ziele. Da diese Abstraktionsebene, wie bei anderen Rollenkonzepten, nur punktuell, dafür aber an entscheidenden Stellen eingenommen werden kann, stünde sie auch nicht im Konflikt mit den zentralen Positionen, die in beiden Ansätzen vertreten werden. Es wäre lediglich ein weiteres Instrument, um die Förderung der in den Ansätzen angestrebte Menschorientierung in Gestaltungsprozessen zu unterstützen.

Die Zuwendung zum Rollenkonzept des Individuums bietet darüber hinaus weitere für CTA und VSD wünschenswerte Effekte, denn durch seine Verwendung wird das potenzielle Risiko minimiert, relevante Bedürfnisse und wahrgenommene Zustände von an Gestaltungsprozessen Beteiligten, besonders den von Technik betroffenen Menschen zu vernachlässigen oder nicht zu berücksichtigen. Die aktive Auseinandersetzung mit den Erfahrungen von Individuen und ihren Perspektiven auf Geschehnisse in Gestaltungsprozessen hält wertvolle und ausschlaggebende Informationen bereit, die in beiden Ansätzen ein umfassenderes Verständnis des Gestaltungsprozesses ermöglichen. Ein möglichst umfassendes Verständnis von Gestaltungsprozessen liegt im Interesse der Positionen beider Ansätze, um zur Gestaltung technischer Systeme beizutragen, die tatsächliche Bedürfnisse und Probleme auf für Menschen angemessene Weise adressieren. Die Auseinandersetzung und Verwendung des Rollenkonzepts des Individuums steht dabei nicht im Widerspruch zu dem Argument, dass technische Systeme in der Regel nicht ausschließlich für einzelne Individuen gestaltet werden. Vielmehr könnte das Begeben auf die niedrigere Abstraktionsebene als weiteres wertvolles Hilfsmittel in Betracht gezogen werden, um vertiefendes Wissen zu Vorbehalten oder negativen Folgewirkungen zu erheben, oder um beispielsweise Wertekonflikte zu antizipieren. Durch die Untersuchung, wie ein technisches System oder eines seiner Features sich auf eine konkrete Lebenssituation eines Individuums auswirkt, kann ermittelt werden, ob sich die dort gezeigten Effekte auch in Lebenssituationen weiterer Individuen und Gruppen äußern. Dadurch können

relevante, da übertragbare Gestaltungsfaktoren identifiziert werden, um nachgelagert entsprechende Förderungs- oder Präventionsstrategien zu formulieren. Und auch wenn dadurch nicht übertragbare Gestaltungsfaktoren zum Vorschein kommen, schließen die Gestaltungsphilosophien einer wertsensitiven und sozialorientierten TG nicht aus, Lösungen für mögliche Randszenarien entwickeln zu können. Gezielt angewendet eröffnet die Betrachtung des Menschen als Individuum daher aufschlussreiches Wissen für Gestaltungsverantwortliche und Entscheidungsträger, das durch Gestaltungsansätze wie CTA und VSD systematisch erschlossen, validiert und in Gestaltungsprozesse integriert werden kann.

### 5.6.3 Risiken des Nutzerrollenkonzepts

In CTA wie auch in VSD wird das Rollenkonzept der Nutzer durchgängig verwendet, um einen Gegenpol zu den Rollenkonzepten der Gestalter zu markieren. In vielen Disziplinen, die sich mit der Gestaltung von Technik auseinandersetzen, ist diese grundlegende Unterscheidung Teil einer in der Vergangenheit etablierten Lingua franca, die in Gestaltungskontexten überwiegend pragmatische Zwecke erfüllen soll. Etwa seit den 2010er-Jahren werden jedoch die mit dem Rollenkonzept verbundenen Implikationen zunehmend kritisch betrachtet, stellenweise auch eine Revision des traditionellen Rollenverständnisses von Nutzern (insbesondere in Kontexten der IT-Gestaltung) eingefordert. Diese Kritik bezieht sich auf die Erkenntnis, dass insbesondere bei der Gestaltung komplexer und vernetzter Systeme die überwiegend linear gedachten Input-Output-Kausalitätsketten nicht länger tragbar sind, die in vielen Ansätzen und Modellierungsverfahren zur Gestaltung digitaler Systeme Anwendung finden. Stark verkürzt beschreiben derartige Modelle häufig Vorgehensweisen, nach denen Gestalter sich vor dem Beginn ihrer aktiven Gestaltungsarbeit einen Überblick über die Eigenschaften und Anforderungen der vermeintlichen Nutzer eines zu gestaltenden Systems erarbeiten, die daraus entstehenden Eindrücke in ein „Abbild der Nutzer“ gießen und dieses ihren nachfolgenden Gestaltungsaktivitäten als Orientierungsgröße mit Repräsentationsanspruch zugrunde legen. Solche Vorgehensweisen bergen jedoch bereits auf konzeptioneller Ebene die Gefahr, dass anstelle eines angemessenen Abbilds menschlichen Lebens und Handelns, das Rollenkonzept der Nutzer ein stark reduziertes „Abziehbild“ darstellt (Wendland, 2021b), welches lediglich die für die es formulierenden Gestalter vertrauten und deshalb genehmen Ausschnitte und Attribute in Betracht zieht. Sämtliche störenden Attribute von Nutzern, die beispielsweise mit bestimmten Weltanschauungen oder Argumentationslinien von Gestaltern oder Projektverantwortlichen kollidieren, oder die Use Cases eines Systems in den Vordergrund rücken, deren Realisierung mit erhöhtem Ressourcenaufwand verbunden wäre, werden durch das Argument nicht vorhandener Relevanz kaum betrachtet oder gänzlich ausgeblendet (Silvast et al., 2018). Das auf diese Weise entstandene Abbild von Nutzern wird so gedanklich wie faktisch schnell beschnitten und „zurechtgerückt“, was eine Entfernung von den tatsächlichen und relevanten Eigenschaften, Bedürfnissen, Verhaltensweisen und Werten von Gruppen und Individuen zur Folge haben kann. Wie sich dies konkret äußern kann und welche Hintergründe damit zusammen hängen, fassen Gonzatto und van Amstel (2022) treffend in einer Publikation zusammen, in der sie die

„Schattenseiten“ des Nutzer-Rollenkonzepts in der Diskurstradition der HCI-Community beleuchten:

*„These people are abstracted from their historical or community references, stereotyped, pathologized, disembodied, and disempowered to become passive beings on the designers' hands, similarly to an imaginary friend. [...] The prevailing conceptualization of so-called users as abstract minds that freely interact with computers ignore the concrete characteristics of their physical bodies, but also the equally concrete characteristics of their social bodies – race/ethnicity, sex/gender, class, disability, and others“* (Gonzatto & van Amstel, 2022, S. 759).

Damit ist im Kern die Problematik der traditionellen Erstellung und Verwendung des Nutzerrollenkonzepts dargelegt, denn es ignoriert häufig kulturanthropologische, ethische oder gesellschaftliche Aspekte, die jedoch in VSD und CTA von besonderer Bedeutung sind. Seine unreflektierte Verwendung wirft insbesondere ethische Fragestellungen auf, da das Nutzerrollenkonzept, wie von Gonzatto und van Amstel beschrieben, dehumanisierende Tendenzen aufweist. Diesen gilt es besonders im Kontext der MZKG eine hohe Aufmerksamkeit zu widmen und sie durch entsprechende Maßnahmen zu entkräften und zu adressieren. So schlägt beispielsweise Rafael (2023) vor, anstelle des Rollenkonzepts Nutzer das Rollenkonzept der „Interactors“ zu etablieren, um zusätzlich zu den funktionalen, technischen und ästhetischen Aspekten, die mit dem Rollenkonzept verbunden sind, auch seinem ethischen Bedeutungsgehalt gerecht werden zu können. Die Autorin verwendet in ihrer Darstellung einerseits Gedanken des Posthumanismus, verweist aber auch auf Don Norman (siehe Kap 2.2.3), der bereits 2006 die Position vertrat, Worte wie *consumer*, *customer* und *user* aus dem Vokabular von Gestaltern zu streichen (Norman, 2006), um den diesen Begriffen anhaftenden dehumanisierenden Tendenzen Einhalt zu gebieten. Denn nicht nur haben diese Tendenzen Auswirkungen auf die Zeit vor und während der Gestaltung technischer Systeme, im äußersten Fall kann auch das Phänomen auftreten, dass das manifestierte krude Abbild von Nutzern in technischen Systemen sich nach deren Markteinführung zu einem neuen Vorbild für menschliches Leben und Handeln entwickelt (Wendland, 2021b).

In der CTA- und VSD-Literatur kann ich die angesprochenen dehumanisierenden Tendenzen des Nutzerrollenkonzepts nicht feststellen, da in beiden Ansätzen viele Bestrebungen vorhanden sind, um genau dieses Risiko zu minimieren. In beiden Ansätzen wird danach gestrebt, die Durchführung von Reflexionsarbeit innerhalb von Gestaltungsprozessen zu fördern. Aufgrund dieser Zielsetzung wird in beiden Ansätzen ein hoher Grad an Selbstreflexion etabliert, der dazu verwendet wird, über die eigenen Standpunkte sowie über diejenigen von Gestaltern und Entscheidungsträgern zu reflektieren. In CTA und VSD wird zudem einen Blick auf Gestaltungsprozesse angewendet, der ihre Dynamik betont und die die mit ihnen verbundene Komplexität wenig reduziert. Auch dies kann im Diskurs um das Rollenkonzept der Nutzer aufschlussreich sein, da die damit verbundenen Positionen einer starren und stark einseitigen Betrachtung von Nutzern entgegenstehen. Die CTA- und VSD-Literatur kann durch ihren mittlerweile über 30 Jahre alten Korpus einen umfangreichen Erfahrungsschatz zum Diskurs beisteuern, der die deutlich jüngere Debatte zur

dehumanisierenden Tendenz des Nutzerrollenkonzepts um aufschlussreiche Fallstudien und Erfahrungsberichte außerhalb der typischen Gestaltungskontexte im IT-Bereich ergänzen kann. Besonders die in VSD verankerten Vorkehrungen zur Erhebung und Berücksichtigung menschlicher Werte, sowie die dort stellenweise durchgeführte Substitution des Rollenkonzepts durch die direkten und indirekten Stakeholder können wegweisend für den künftigen Umgang mit dem Rollenkonzept Nutzer sein – für benachbarte Felder wie auch in CTA und VSD.

#### 5.6.4 Fehlende Betrachtung von Managementrollenkonzepten

In beiden Ansätzen wird viel Mühe darauf verwendet, sämtliche Rollenkonzepte zu erfassen und zu beschreiben, die für die Anstellung von Betrachtungen auf den Abstraktionsebenen in CTA und VSD relevant sind. Wesentlicher Bestandteil dieser Anstrengungen ist es, die verschiedenen Zuständigkeitsbereiche der Rollenkonzepte darzustellen und voneinander abzugrenzen. Entsprechend stehen im Zentrum vieler Betrachtungen die verschiedenen Sozial- und Wertefragen berührenden Dynamiken, die nach Ansicht in CTA und VSD deutlich mehr Raum innerhalb von Entscheidungsprozessen einnehmen sollten. Bezüglich der Frage, welchen Rollen zugeschrieben wird, die wichtigen und zentralen Gestaltungsentscheidungen zu treffen, vertreten CTA und VSD ähnliche Positionen. So wird in der CTA-Literatur deutlich, dass diese Entscheidungsmacht häufig auf Seiten politischer Entscheidungsträger oder bei Akteuren des Technologiefelds verortet wird, die die wirtschaftlichen Ressourcen zur Entwicklung einer Technologie beisteuern (two-track-approach). In VSD wird diese Entscheidungsmacht den Gestaltern zugeordnet, die für die technische und konzeptuelle Struktur technischer Systeme verantwortlich sind. Abgesehen von den politischen Entscheidungsträgern in CTA wird in beiden Ansätzen die Verantwortung für Gestaltungsentscheidungen demnach bei den Gruppen und Individuen verortet, die Technik konzeptuell und faktisch gestalten. Eine derartige Verortung von Gestaltungsmacht beinhaltet jedoch einen blinden Fleck, denn in beiden Ansätzen wird bei den Rollenkonzepten der Gestalter nicht zwischen technischen und organisatorischen Management-Fragestellungen unterschieden. In CTA wie auch in VSD wird durch die Verwendung der Rollenkonzepte der Gestalter häufig implizit angedeutet, beide Facetten der technischen und organisatorischen Gestaltungsverantwortung zu bezeichnen. Während Gestalter in der Tat für viele Aspekte der Gestaltung Verantwortung tragen, sind in vielen Gestaltungsprozessen hingegen Menschen mit Aufgaben beauftragt, die sich ausschließlich Fragestellungen zur organisatorischen Herausforderungen von Gestaltungsprozessen widmen. Besonders in Gestaltungsprozessen, in denen komplexe und vernetzte Systeme gestaltet werden, erfährt die Unterscheidung der technischen und organisatorischen Verantwortung einen zunehmend bedeutenden Stellenwert. Häufig sind mit den mehr organisatorischen Aufgabenprofilen Rollenkonzepte wie Projekt- oder Produktmanager assoziiert, die beispielsweise zur Aufgabe haben, den Fluss von Ressourcen zu planen, zu koordinieren und zu kontrollieren, sowie den Gesamtfortschritt von Gestaltungsprozessen zu überwachen. Sie stellen Gestalter mit sämtlichen Ressourcen, Werkzeugen und Materialien aus, die sie für eine zielführende Durchführung ihrer Gestaltungsarbeit benötigen. Aufgrund ihrer Position an entscheidenden Schnittstellen sowie ihrer

Absicherungs- und Kontrollfunktion verfügen Managerrollen entsprechend häufig über einen beträchtlichen Anteil der Entscheidungsmacht, die nicht zuletzt aufgrund der Verwaltung personeller und finanzieller Ressourcen zustande kommt. Häufig verfügen sie deshalb auch über deutlich mehr Entscheidungsmacht, als die in VSD und CTA häufig zur Verantwortung gezogenen Gestalter.

Insofern ist es überraschend, dass diesen Rollenkonzepten innerhalb der VSD- und CTA-Literatur vergleichsweise wenig Aufmerksamkeit gewidmet wird. In beiden Ansätzen wird die Zielsetzung verfolgt, Entscheidungsprozesse an entscheidenden Stellen positiv zu beeinflussen, jedoch setzt sich die Literatur dafür kaum mit den Managementrollenkonzepten auseinander. Dies ist auch beim Blick auf die Ursprünge der Rollenkonzepte verwunderlich, da in beiden Ansätzen auf verschiedentliche Erkenntnisse der Fachbereiche der Ökonomie und der Managementlehre Bezug genommen wird, um eigene Positionen auszuarbeiten und argumentativ zu unterstützen. Mit Rollenkonzepten ist häufig eine Abgrenzungsfunktion verbunden, die häufig das Eingehen von Kompromissen beinhaltet. So beschreiben auch Friedman und Hendry (2019, S. 43) an einer Stelle, dass sich die Limitierung der betrachteten Stakeholdergruppen in VSD häufig als eine zentrale Herausforderung in VSD-Projekten darstellt. Dies lässt den Schluss zu, dass zumindest in den konzeptionellen Ausarbeitungen zu VSD das Fehlen der Managementrollenkonzepte ein Resultat dieser notwendigen Kompromissfindung darstellt. Mit Blick auf die Gestaltung komplexer technischer Systeme wie in der KI-Gestaltung sollte in beiden Ansätzen künftig jedoch eine stärkere Betrachtung von Managementrollenkonzepten erfolgen, um dem selbst formulierten Anspruch einer ganzheitlichen Betrachtung von Gestaltungsprozessen gerecht zu werden.

### 5.6.5 Zum Auswirkungsanspruch von CTA agents

In den Erläuterungen zum Rollenkonzept der CTA agents werden in der CTA-Literatur einige Anstrengungen unternommen, um darzulegen, dass es sich bei ihnen *nicht* um ein Gestalter-Rollenkonzept handelt. Sie unterscheiden sich von Gestaltern dadurch, dass sie keine „Hand an die Technik legen“, dass sie keine technischen Funktionen definieren oder realisieren. Auch können sie nicht zu anderen Entscheidungen herbeiführenden Rollenkonzepten gezählt werden, da sie an den wegweisenden Entscheidungsprozessen nicht beteiligt sind und in diesen kein Stimmrecht besitzen oder ein solches einfordern. CTA agents sollen Einfluss auf den Verlauf von Gestaltungsprozessen nehmen, jedoch soll dieser „sanft“ und „indirekt“ erfolgen, indem CTA agents Funktionen ähnlich jenen von Moderatoren oder Vermittlern bekleiden. Ein zentrales Attribut von CTA agents ist gemäß den Interpretationen der CTA-Literatur ein hohes Maß an Neutralität und Objektivität, die sie sich durch die beständige Wahrung von Distanz zu richtungsweisenden Fragen und Entscheidungen von Gestaltungsprozessen bewahren. Diese Auslegung des Rollenkonzepts weist daher insofern einen blinden Fleck auf, da es sich nicht mit der Diskrepanz auseinandersetzt, die zwischen den von CTA agents durchgeführten Tätigkeiten, und der mit diesen Tätigkeiten in Anspruch genommenen Bedeutsamkeit für Gestaltungsprozesse besteht. Die konstant aufrechterhaltene Distanz und die bewusste Vermeidung einer direkten Beteiligung an Entscheidungsprozessen läuft dem mit dem

Rollenkonzept beanspruchten Verantwortungskontingent und seiner Einflussreichweite entgegen. Der (selbst definierte) Zuständigkeitsbereich der CTA agents erstreckt sich lediglich auf die Durchführung der Interventionsmethoden und endet mit dem Ende eines CTA-Workshops. Von jeder darüber hinaus reichenden Verantwortung nimmt die CTA-Literatur Abstand. In diesen Workshops, die häufig nur wenige Tage in mehreren Jahren umfassenden Technologieentwicklungsprozessen andauern, streben CTA agents danach, durch bewusst gesetzte Impulse Veränderungen im Denken und Handeln der Technologiefeldakteure herbeizuführen. Im Interesse von CTA agents liegt es, Reflexion und Einsicht zu fördern, sodass die künftige Entwicklung einer Technologie aus einer von den Technologiefeldakteuren stammenden intrinsischen Motivation heraus erfolgt. Die Literatur nimmt in Anspruch, dass den CTA agents dies gelingt, stellt jedoch kaum Hilfsmittel bereit, um diese Hypothese zu überprüfen und die Erfolge der Interventionsmaßnahmen zu validieren. Die Darstellungen zu den von CTA agents herbeigeführten und richtungsweisenden Veränderungen in Gestaltungsprozessen wirken deshalb stellenweise überzeichnet, da das Beharren auf Distanzwahrung sich konträr zu dem notwendigen Grad an Beteiligung verhält, die für den in Anspruch genommenen „impact“ erforderlich wäre. So kann bezüglich der Bedeutsamkeit der Aktivitäten von CTA agents der irritierende Eindruck eines „cherry pickings“ entstehen, der Relevanz in Anspruch nimmt, jedoch davor zurückschreckt, die dafür notwendige Verantwortung zu übernehmen und keine Bereitschaft signalisiert, größere Verpflichtungen einzugehen.

Auch beim Blick auf die Aktivitäten, die CTA agents durchführen, wird der Aspekt einer fehlenden Verbindlichkeit deutlich. Die CTA-Literatur nimmt in Anspruch, dass die von CTA agents angewandten Interventionsmethoden positive Effekte auf Gestaltungsprozesse haben. Eine entscheidende Voraussetzung für das Eintreten jener positiven Effekte wäre jedoch, dass als ein Ergebnis von CTA-Workshops die Workshopteilnehmer eine Vereinbarung herbeiführen, die die Übernahme von verbindlichen Verpflichtungen für das weitere Gestaltungsgeschehen beinhaltet:

*„the workshop seemed to have provided a forum for lively discussion and a number of innovative ideas. None of these have been implemented yet, however, at the time of writing this article“* (Heiskanen, 2005, S. 66).

Ein solches Element sieht jedoch die Aufgabenbeschreibung der CTA agents in der Literatur nicht vor. Die Teilnahme an den Workshops sowie das Bereitstellen von Informationen in den Workshops erfolgt für Teilnehmer auf vollständig freiwilliger Basis. Wohl beansprucht die Literatur als Attribut des Rollenkonzepts, über eine Expertise im Leisten von Überzeugungsarbeit für eine Teilnahme zu verfügen und eine Umsetzung von Ergebnissen bei Technologiefeldakteuren gewährleisten zu können. Jedoch gewährt die Literatur kaum Einblicke dahingehend, wie eine solche Argumentationsarbeit inhaltlich strukturiert werden kann. So finden sich beispielsweise kaum Hinweise darauf, welche Maßnahmen CTA agents ergreifen können, wenn ihre Argumente kein Gehör finden. Ein weiterer Aspekt, der die Limitierungen des Handlungsspielraums von CTA agents aufzeigt, besteht in dem Mangel an Ausführungen zum Umgang mit den Ergebnissen und Erkenntnissen der Interventionsmaßnahmen. Da CTA agents nicht zu den faktischen Entscheidungs-

trägern zählen, verbleibt die Entscheidung zum Umgang mit den Workshopergebnissen der CTA-Workshops zu großen Teilen bei den Workshopteilnehmern. So besteht durchaus die Möglichkeit, dass sich die Teilnehmer dazu entscheiden, die Workshopergebnisse für den weiteren Verlauf eines Gestaltungsprozesses gänzlich zu ignorieren, wodurch die CTA-Interventionsmaßnahmen keinerlei Wirkung entfalten könnten. Wie auch die Teilnahme an den Workshops, erfolgt die Implementierung von Veränderungen durch die Teilnehmer auf freiwilliger Basis. Die CTA-Literatur gibt CTA agents keine Werkzeuge an die Hand, die sie dazu befähigen würden, die Workshopteilnehmer stärker zur Umsetzung von Veränderungsmaßnahmen zu motivieren, die sich auf die in den Workshops identifizierten, wünschenswerten Effekte beziehen. Eine Umsetzung von Maßnahmen basiert ausschließlich auf dem Erkennen ihres positiven Wertes für die Gesellschaft durch die Workshopteilnehmer. Die Entwicklung dieser Erkenntnis nimmt die CTA-Literatur dabei häufig als gegeben an.

Da mit dem Ende eines CTA-Workshops die Zuständigkeit von CTA agents im aktiven Gestaltungsgeschehen endet, betrachtet die CTA-Literatur eine entscheidende Stelle im Prozess – das Zeitfenster *nach* der Durchführung von CTA-Workshops – in überraschend geringem Umfang. CTA agents sollen sich nach der Durchführung der Workshops wieder aus dem Technologiefeld und somit dem aktiven Gestaltungsgeschehen zurückziehen, das nachfolgende Gestaltungsgeschehen beobachten und die positiven Entwicklungen identifizieren, die sich auf ihre Interventionstätigkeit zurückführen lassen. Die Beobachtung negativer Entwicklungen wird entsprechend als Anlass betrachtet, erneute Interventionsmaßnahmen anzuberaumen. Da CTA agents überwiegend eine Repräsentation von TA-Arbeit verrichtenden Gruppen und Individuen darstellen, erfolgt die Auseinandersetzung mit der Zeit nach den Workshops sowie mit den Workshopergebnissen in der CTA-Literatur häufig auf wissenschaftlicher Ebene vor dem Hintergrund des wissenschaftlichen Erkenntnisgewinns.



Die Herbeiführung tatsächlicher Veränderungen in Gestaltungsprozessen wirkt im Vergleich dazu, entgegen den Beschreibungen des Rollenkonzepts, häufig mehr wie ein sekundäres Ziel. Dadurch kann das Argument angebracht werden, dass trotz dem in CTA zum Ausdruck gebrachten Plädoyer für eine Distanzminimierung von TA zu TG (siehe Kap. 4.2), CTA agents ihre TA-Komfortzone in CTA nur geringfügig erweitern müssen.

### 5.6.6 Differenzierung der Motivationen von Enactors

Die Darlegung der Ursprünge der Rollenkonzepte der Enactors und comparative Selectors (Kap. 5.5.2) hat gezeigt, dass sie sich in ihrer ursprünglichen Form auf Gruppen und Individuen beziehen, die sich in wissenschaftlichen Professionsbereichen mit TG auseinandersetzen. Der Text von Garud und Ahlstrom, dem die Rollenkonzepte entnommen sind, stammt aus der Managementforschung und strebt danach, die besonderen Verhältnisse und Dynamiken von in TG-Prozessen aktiven Insidern im Verhältnis zu den mehr von den Prozessen entfernten und auf technischer Ebene weniger versierten Outsidern zu ergründen. Die Intention der beiden Autoren besteht jedoch *nicht* darin, Einfluss auf Gestaltungsprozesse zu nehmen, sondern lediglich eine bestimmte, in *Forschungskontexten* anzutreffende soziale Konstellation zu betrachten und Verbindungen zu Entscheidungsprozessen auf politischer Ebene herzustellen. Für die Betrachtungen, die Garud und Ahlstrom anstellen, ist dabei die Unterscheidung von Gruppen und Individuen, die in der Wissenschaft verortet sind, und jenen, die sich außerhalb wissenschaftlicher Organisationen an der Gestaltung technischer Systeme beteiligen, kaum von Bedeutung. Der Transfer der Rollenkonzepte der Enactors und comparative Selectors auf Organisationsformen in Politik und Wirtschaft erfolgt erst in CTA. In CTA werden die ursprünglich für Wissenschaftler formulierten Rollenkonzepte insbesondere mit Gestalterrollenkonzepten gleichgesetzt, die Theorie von Garud und Ahlstrom wird auf einen neuen Anwendungsbereich ausgedehnt. Dabei wird in CTA kaum eine Unterscheidung von Gestaltern in wissenschaftlichen Tätigkeitsfeldern und Gestaltern in wirtschaftlichen Tätigkeitsfeldern durchgeführt, da diese für Betrachtungen auf der Abstraktionsebene, die in CTA Anwendung findet, nur eine geringe Relevanz besitzen und deshalb im CTA-Denken vernachlässigt werden. Jedoch sind mit dieser Gleichstellung Implikationen verbunden, die eine nähere Betrachtung erfordern.

So wird in CTA suggeriert, dass Gestalter innerhalb und außerhalb wissenschaftlicher Einrichtungen als Synonym betrachtet und verwendet werden können, da sie sich die entscheidende Eigenschaft teilen, von Technik überzeugt und begeistert zu sein. Dabei wird jedoch außer Acht gelassen, dass ein Gestalter, der sich in einem Forschungskontext mit der Gestaltung eines technischen Systems befasst, fundamental andere Motivationen besitzen kann, wie beispielsweise ein Gestalter, der den Auftrag hat, ein technisches System zu gestalten, das primär zum Ausbau der Marktposition eines Unternehmens beitragen soll. Den beiden Rollenkonzepten wird in CTA demnach mit Verweis auf das Begeisterungsargument eine Vergleichbarkeit zugeschrieben, die in Technologiefeldern nicht zwingend vorhanden sein muss, und die in CTA-Projekten bei Feldakteuren Fragen und Irritationen aufwerfen kann. Die Prämisse dieser Gleichstellung wirkt auch insofern überraschend, da der CTA-Literatur mit der Nanotechnologie ein Technologiefeld als

Anschauungs- und Erprobungsbeispiel dient, in dem eben jene Unterschiede von mehr wissenschaftlich motivierten und mehr wirtschaftlich motivierten Gestaltern – trotz ihrer geteilten Begeisterung – zum Ausdruck kommt. Da sich die CTA-Literatur viel Mühe gibt darzustellen, die Dynamiken in Technologiefeldern möglichst korrekt zu erfassen und in den CTA-Workshops wiederzugeben, kann es für eine Weiterentwicklung des Ansatzes daher hilfreich sein, im Sinne einer Selbstreflexion die Prämisse der Gleichstellung zu überdenken und auf Detailebene weiter auszuarbeiten. Es bestehen Möglichkeiten Rollenkonzepte zu formulieren, die das in CTA entscheidende Attribut der Enactors, ihre Unzugänglichkeit für mögliche Bedenken und Gegenargumente technischer Lösungen, zum Ausdruck zu bringen, und dabei dennoch die möglichen Unterschiede dieser Position zugrundeliegenden Motivationen hervorzuheben. In Summe wird durch die Rollenkonzepte der Enactors und comparative Selectors deutlich, dass CTA trotz vieler Bemühungen ein Ansatz von Wissenschaftlern für Wissenschaftler ist, die sich selbst primär auch als solche verstehen – als Wissenschaftler, nicht als Gestalter. Dies mindert die Anschlussfähigkeit von CTA für stark auf die praktische Gestaltung ausgerichtete Diskurse wie den MZKG-Diskurs in HCAI.

### 5.6.7 Unvollständige Auseinandersetzung mit Technikgegnern

In der Literatur zu beiden Ansätzen kommt zum Ausdruck, dass in TG-Prozessen neben Technikbefürwortern häufig auch Technikgegner vertreten sind, die zu Gestaltungsprozessen valide Argumente beisteuern. Ihnen soll deshalb eine angemessen hohe Aufmerksamkeit gewidmet werden und die Berücksichtigung ihrer Positionen sollte sich im Produkt von Gestaltungsprozessen wiederfinden. Technikgegner sind besonders in CTA ein relevantes Rollenkonzept, da eine selbst erlegte Aufgabe in CTA darin besteht, die von Technik begeisterten Enactors mit den validen Argumenten der Opponents zu konfrontieren, um sie zur Reflexion zu bewegen und etwas in der Ausführung ihrer „Mission“ zur Ausräumung von Hindernissen auszubremsen. Auch die VSD-Literatur setzt sich mit Technikgegnern insofern auseinander, dass sie dort häufig zentrale Figuren in Wertekonflikten darstellen, deren Positionen und die damit verbundenen Hintergründe es in VSD-Projekten zu ermitteln gilt. In beiden Ansätzen scheint entsprechendes Wissen zu Technikgegnern vorhanden zu sein, jedoch werden Technikgegner in der Literatur überwiegend auf einer Argumentationsebene als Contra-Seite konzipiert und dargestellt. In Gestaltungsprozessen treten Technikgegner jedoch nicht nur verbal, sondern auch aktiv handelnd in Erscheinung. Sie lassen ihren Worten in Rollen, die als „Störer“ oder „Verhinderer“ bezeichnet werden könnten, Taten folgen, um den Fortschritt von Gestaltungsprozessen zu behindern, zu sabotieren oder in gänzlich andere Bahnen zu lenken. Insofern besteht der blinde Fleck der Rollenkonzepte in CTA und VSD darin, dass Technikgegner nicht in ihrem vollständig relevanten Attributskorpus betrachtet werden. Hierin liegt die Gefahr einer Unterschätzung, denn gerade die tatsächlichen Handlungen von Technikgegnern können große Auswirkungen auf Gestaltungsprozesse haben, die es vorausschauend zu erkennen und möglichst abzuwenden gilt. Dieser Umstand deckt sich mit den in beiden Ansätzen zum Ausdruck gebrachten Intentionen und zumindest in der CTA-Literatur scheinen die damit verbundenen potentiellen Auswirkungen bekannt, da dort an einer

Stelle das historische Beispiel der Ludditen referenziert wird (Schot & Rip, 1997). In CTA und VSD wird jedoch nach derzeitigem Kenntnisstand nicht die Chance wahrgenommen, im Rahmen der Auseinandersetzungen mit Technikgegnern hervorzuheben, dass zusätzlich zu ihrer Identifikation und dem Gehörschenken ihrer Positionen von zentraler Bedeutung ist, sie wertschätzend und zielführend in Gestaltungsprozesse zu integrieren. In der Erarbeitung in entsprechenden Methoden und Strategien läge ein großer Mehrwert, der beiden Ansätze ein Alleinstellungsmerkmal bescheren und darüber hinaus ihre Relevanz in übergreifenden Gestaltungsdiskursen wie dem zur MZKG steigern würde.

## 5.7 Diskussion

Die Untersuchung der Rollenkonzepte in CTA und VSD hat gezeigt, welche Annahmen zu menschlichem Leben und Handeln in beiden Ansätzen zugrunde gelegt werden. Es wurde deutlich, dass in beiden Ansätzen unterschiedliche Abstraktionsebenen gewählt werden, um die verschiedenen Formen und hohen Grade von technischer, organisatorischer und sozialer Komplexität von Gestaltungsprozessen zu ermitteln und ihnen zu begegnen. In CTA wird der Mensch vorwiegend als Akteur betrachtet, der in verschiedenen sozialen Settings Handlungen durchführt, um zur Gestaltung technischer Systeme beizutragen. Handlungen und Entscheidungen in Gestaltungsprozessen werden dabei aus einer Position betrachtet, die primär die Bedeutungen und Auswirkungen der Handlungen für eine Gesellschaft erfragt und antizipiert. Entsprechend finden in CTA recht hohe Abstraktionsebenen Verwendung, da nur mit ihrer Hilfe eine Extrapolation auf weite Zeiträume und für eine hohe Anzahl an Menschenleben erfolgen kann. In VSD wird der Mensch vorwiegend als Stakeholder betrachtet, der spezifische Motivationen und Interessen besitzt, die ihn zur Gestaltung oder zur Nutzung von technischen Systemen bewegen. Handlungen und Entscheidungen werden überwiegend auf einer Projektebene untersucht, dessen Auswirkungen auf angrenzende Gesellschaftsteile im Rahmen eines wertsensitiven Gestaltungsprozesses sorgsam überprüft werden soll. Entsprechend findet in VSD eine höhere Varianz an Abstraktionsebenen Verwendung, die verschiedene Stufen an Extrapolation oder Detailbetrachtung menschlichen Lebens und Handelns erlauben (siehe Abb. 43).



Abbildung 43: Abstraktionsebenen, mit deren Hilfe in VSD menschliches Leben und Handeln untersucht wird (B. Friedman & Hendry, 2019, S. 21)

Wie in CTA und VSD lautet auch in HCAI der Anspruch, durch die Identifikation und Berücksichtigung bedeutsamer Faktoren menschlichen Lebens und Handelns Gestaltungserfolge herbeizuführen. Anders als in CTA und VSD steht in HCAI jedoch die KI-Technologie mit ihren spezifischen Eigenschaften und Herausforderungen im Vorder-

grund. Ferner beziehen sich die in HCAI vertretenen Positionen der namensgebenden Menschzentrierung und die dazu angewandten Abstraktionsebenen auf die in der MZG formulierten Gestaltungsprinzipien (siehe Kap. 2.3). Die Umsetzung dieser Gestaltungsprinzipien trifft im Kontext der MZKG vermehrt auf geistes- und sozialwissenschaftliche Fragestellungen, die für in HCAI typische Denkweisen neue Herausforderungen darstellen (siehe Kap. 3.3). Ein angemessener Umgang mit der höheren sozialen, ethischen, technischen und organisatorischen Komplexität von KI erfordert eine Erweiterung der HCAI-Gestaltungsphilosophie, insbesondere um die Möglichkeit, temporär höhere Abstraktionsebenen einnehmen zu können, um menschliches Leben und Handeln zu betrachten und zu untersuchen. Ein Blick auf die Rollenkonzepte in CTA und VSD kann Impulse liefern, um die Entwicklung von in HCAI benötigten Strategien zur Verbindung von Menschorientierung und Gestaltungspragmatik zu informieren und zu inspirieren. Inwieweit die Rollenkonzepte hierfür eine Hilfestellung darstellen können, diskutiere ich nachfolgend, indem ich die Erkenntnisse zu Bedeutungen, Zwecken, Ursprüngen und blinden Flecken der Rollenkonzepte aufgreife.

### 5.7.1 Erweiterung von Denkhaltungen und Perspektiven

Wie die vergleichende Analyse gezeigt hat, wird der Reflexion verwendeter Rollenkonzepte in CTA wie auch in VSD ein hohes Maß an Aufmerksamkeit gewidmet. Stellenweise gibt die Literatur Detailinformationen zu einzelnen Rollenkonzepten nur in geringem Umfang preis, doch trotz dieser stellenweise vorhandenen Vagheit ist in der Literatur zu beiden Ansätzen ein grundsätzlich hohes Bewusstsein zur Bedeutsamkeit und den verschiedenen Funktionen von Rollenkonzepten erkennbar. In CTA wird dies beispielsweise bei Schot und Rip deutlich, wo das Rollenkonzept der CTA agents vorgestellt wird und die beiden Autoren eine Reflexion über das mit CTA agents verbundene Werteprofil anstellen (Schot & Rip, 1997, S. 265). Die beiden Autoren demonstrieren ferner eine grundsätzliche Reflexionsfähigkeit zu Rollenkonzepten, da sie die Notwendigkeit aufzeigen, sämtliche relevanten Akteursrollen zu identifizieren und sie im Rahmen der CTA-Workshops zusammenzuführen (Schot & Rip, 1997, S. 257). Auch in der VSD-Literatur wird deutlich, dass über die Verwendung von Rollenkonzepten und den damit verbundenen Implikationen reflektiert wird und dass die in VSD verwendeten Rollenkonzepte bewusst ausgewählt und verwendet werden. So heben beispielsweise Davis und Nathan im Zusammenhang mit der Methode der Stakeholderanalyse hervor, dass der Begriff Stakeholder sich auf Rollen bezieht, die je nach betrachtetem Gestaltungskontext unterschiedliche Bedeutungen besitzen:

*„Note that the term stakeholders refers to roles and not individual people. [...] Conceptualizing an individual as mother rather than as in the role of mother risks ignoring the multiplicity of roles through which we engage with our environments“ (Davis & Nathan, 2015, S. 17).*

Dies zeigt, dass auch in VSD die Pluralität vorhandener Rollen in Gestaltungskontexten anerkannt wird, und dass die Verwendung der Rollenkonzepte in VSD einen bewusst gewählten Bruchteil der Rollenvielfalt darstellt, die Menschen besitzen können – jenen

Bruchteil, der Aufgrund seines Bezuges zur interactional stance in VSD Relevanz besitzt. Dies bringen auch Friedman und Hendry an anderer Stelle explizit zum Ausdruck:

*„Stakeholders within value sensitive design are defined by and understood in relationship to their interaction with a technology or sociotechnical system. That is, stakeholders are considered by role, rather than by “person” or other “entity.” A “role” pertains to a stakeholder’s duties, contextual identity, or particular circumstances”* (B. Friedman & Hendry, 2019, S. 40).

In beiden Ansätzen ist somit ein differenziertes Verständnis von Rollenkonzepten vorhanden, das die Notwendigkeit einer detaillierten Auseinandersetzung mit unterschiedlichen Rollenkonzepten und den daraus resultierenden Wert für Gestaltungsprozesse anerkennt. Begrüßenswert ist die Explikation dieses Bewusstseins insbesondere unter dem Gesichtspunkt, dass Rollenkonzepte lediglich eine zugunsten von Gestaltungsprozessen reduzierte Betrachtung menschlichen Lebens und Handelns darstellen, die je nach Gestaltungskontext auf Brennpunkte ausgerichtet, detailliert ausgearbeitet, und somit ausgeweitet werden müssen. Diese grundsätzliche Awareness, die sorgsame Auseinandersetzung und die aus ihr resultierende reflektierte Verwendung von Rollenkonzepten stellt auch für HCAI eine Voraussetzung dar, die vor dem Hintergrund der Zielsetzung einer menschenzentrierten Gestaltung erfüllt werden sollte. Insbesondere in der Auseinandersetzung mit Rollenkonzepten wie den informierten Einflussnehmern liegt dabei ein Mehrwert für HCAI, da sie über die traditionelle Unterscheidung von Gestaltern und Nutzern hinausgehen und dabei unterstützen, die möglichen Bedeutungen von Menschzentriertheit im KI-Kontext zu schärfen.

Der hohe Reflexionsgrad zu Rollenkonzepten ist in VSD und CTA ein Resultat der formulierten Zielsetzung, in Gestaltungsprozessen eine stärkere Reflexion über Motivationen und Werte der gestaltenden und betroffenen Gruppen und Individuen zu fördern. Die dafür ausgeprägte Reflexionsfähigkeit wird in CTA und VSD auch zur Reflexion der eigenen Gedankengänge und Konzepte verwendet, was in den verschiedenen Abstraktionsebenen und dem stellenweise hohen Detailgrad der ausgearbeiteten Rollenkonzepte mündet. Das daraus resultierende, grundsätzlich sehr hohe Bewusstsein zu den Bedeutungen, Funktionen, Implikationen und Limitierungen von Rollenkonzepten ist auch für die in der MZKG-Debatte angetroffenen Herausforderungen aufschlussreich. So stellt auch in HCAI die hohe Komplexität, Dynamik und die daraus resultierende Wandelbarkeit von Rollen im Gestaltungsgeschehen einen zentralen Sachverhalt dar, dem häufig pragmatisch begegnet werden muss. Besonders die mehr unspezifischen und deshalb flexiblen Rollenkonzepte wie beispielsweise der Involvierten können zum Umgang mit diesen Herausforderungen besonders in frühen Gestaltungsphasen neue Perspektiven aufzeigen und neue Möglichkeiten eröffnen, auch wenn in HCAI-Projekten zu fortgeschrittenen Zeitpunkten ihre Offenheit zugunsten konkreter Attribute sicherlich reduziert werden muss. In HCAI kann durch die Verwendung der mehr unspezifischen Rollenkonzepte von CTA und VSD das Bewegen auf und hin zu gedanklich höheren Abstraktionsebenen erlernt und aufgegriffen werden. Dies wird jedoch stellenweise auf Kosten der Genauigkeit von Rollenkonzepten erfolgen müssen, wodurch auch bei der Formulierung und

Verwendung entsprechender Rollenkonzepte in HCAI das Eingehen von Kompromissen erforderlich sein wird. Erfolgt das Eingehen dieser Kompromisse jedoch auf ähnlich reflektierte Weise wie in CTA und VSD, kann eine Wahrung und Förderung der Menschorientierung auch auf den notwendig höheren Abstraktionsebenen gewährleistet werden. So kann die Offenheit der Rollenkonzepte der Involvierten oder der informierten Einflussnehmer in Gestaltungsprozessen einen gedanklichen Raum eröffnen, der es erlaubt, bedeutsame Zusammenhänge und Abhängigkeiten im sozialen Geschehen zu identifizieren, um daraus relevante Gestaltungsfaktoren und Gestaltungspfade abzuleiten. Ferner können die Rollenkonzepte dazu beitragen, diesen gedanklichen Raum über längere Zeiträume geöffnet zu halten und so die Möglichkeit eröffnen, in Gestaltungsprozessen in HCAI länger über Gestaltungsalternativen nachzudenken. Sie beugen einer zu schnellen Eingrenzung auf bestimmte Anwendungsformen oder Entwicklungspfade vor, die aufgrund der höheren sozialen und ethischen Implikationen von KI-Systemen im MZKG-Diskurs oftmals wenig zielführend sind. Die Rollenkonzepte können ferner dazu beitragen, die Ränder der betrachteten Ausschnitte sozialer Realität in HCAI sukzessive auszudehnen und ein Erkennen von und ein Denken in Netzwerkstrukturen etablieren, die in VSD, stärker noch in CTA angewendet werden. Die Förderung und Adaption einer Betrachtung sozialen Geschehens in Netzwerkstrukturen kann in HCAI zusätzlich erleichtert werden, indem die Verbindungen der Rollenkonzepte zur Actor-Network-Theory betrachtet werden (siehe Kap. 5.5.1).

Damit ist ein weiterer potenzieller Mehrwert für die HCAI-Community verknüpft, denn die Verwendung von Rollenkonzepten verschieden ausgeprägter Detailgrade ermöglicht es darüber hinaus, zu verschiedenen Zeitpunkten in Gestaltungsprozessen zwischen verschiedenen Abstraktionsebenen Perspektivwechsel durchzuführen (siehe Kap. 5.7.8). So können durch Rollenkonzepte wie den Technikenthusiasten oder Technikgegnern beispielsweise auf konkreter Ebene Wertefragen oder Wertekonflikte untersucht werden. Ein anschließend bewusst durchgeführter Perspektivwechsel auf dieselben Wertefragen durch beispielsweise Rollenkonzepte der informierten Einflussnehmer kann wertvolle Hinweise zum „bigger picture“ gesellschaftlicher Zusammenhänge eröffnen. Die Rollenkonzepte von CTA und VSD stellen ein Angebot dar, menschliches Handeln in einem weiten Spektrum zu betrachten und anlassbezogen gedanklich hinein- und heraus zu zoomen. In HCAI kann dadurch der geschaffene gedankliche Raum um mehrere Dimensionen erweitert werden, die zur zielführenden Auseinandersetzung mit den durch die KI-Gestaltung aufgeworfenen sozialen und ethischen Fragestellungen benötigt werden. Diesbezüglich kann jedes Rollenkonzept aus CTA und VSD, das im Vergleich mit den traditionellen Denkweisen in HCAI auf den ersten Blick „fremd“ wirkt, hilfreich sein, um für die Gestaltung relevante Aspekte aufzudecken. Zu beachten ist hierbei jedoch, dass die Auseinandersetzung mit diesen Rollenkonzepten in HCAI – anders wie stellenweise in CTA – immer vor dem Hintergrund des Mehrwerts für die tatsächliche Gestaltung erfolgen muss. Der Spielraum für rein technikphilosophische Diskurse im Sinne eines „Nachdenkens über Technik“ (Hubig et al., 2013) ist in HCAI eng bemessen, da solche Aktivitäten die Herbeiführung konkreter Gestaltungsentscheidungen nur in geringem Maße unterstützen. Somit werden die Räume zur Reflexion in HCAI tendenziell schneller und stärker thematisch orientiert werden müssen.

Eine Frage, die auch beim Blick auf die Rollenkonzepte in CTA und VSD offenbleibt ist, anhand welcher Kriterien die Begrenzung der Betrachtung menschlichen Lebens und Handelns in HCAI erfolgen kann bzw. sollte. Besonders die VSD-Literatur geht offen mit dem Umstand um, dass das Setzen von Grenzen in der Betrachtung von Ausschnitten sozialer Realität immer einen Kompromiss darstellt, mit dessen Implikationen sich ein Gestaltungsansatz mit Ansprüchen zur Werteorientierung kontinuierlich auseinandersetzen sollte. In CTA werden hingegen nur in geringem Umfang Betrachtungs- oder Relevanzgrenzen gesetzt, was auf Kosten von Spezifität erfolgt und eine rasche Übertragung und Anwendung auf konkrete Gestaltungssettings tendenziell erschwert. Die Folgewirkungen, die KI-Systeme auf das gesellschaftliche Leben haben können, sind in der Regel schwer abzuschätzen und bewegen sich häufig in sehr weitreichenden Dimensionen. Insofern kann auch in HCAI argumentiert werden, dass die gewählten Grenzen der betrachteten sozialen Ausschnitte tendenziell eher weit gefasst werden sollten. Dies kollidiert jedoch mit der Notwendigkeit der Herbeiführung von Gestaltungsentscheidungen, sodass letztlich auch in HCAI Kompromisslösungen zur Eingrenzung gefunden werden müssen. Zur Herbeiführung dieser können CTA und VSD wenig beitragen, da sie sich nicht ausreichend detailliert mit den Anforderungen der KI-Gestaltung auseinandersetzen. Woran sich HCAI jedoch bei der Identifikation von Grenzpunkten des eigenen Verantwortungsbereichs orientieren kann, ist die transparente Darstellung der dieser zugrundeliegenden Argumente und Gedanken, wie sie in VSD erfolgt. Besonders die Betrachtung der Interaktionsebene von Mensch-Technik-Verhältnissen in VSD kann hierbei eine geeignete Orientierung darstellen, da diese häufig auch in den Debatten zur MZKG von zentraler Bedeutung ist.

### 5.7.2 Interpretationen identisch bezeichneter Rollenkonzepte

Ein weiterer Anknüpfungspunkt von HCAI zu CTA und VSD besteht in der sprachlichen Deckungsgleichheit einiger Rollenkonzepte, die in den drei Gestaltungsfeldern verwendet werden. So fällt auf, dass auch in HCAI Rollenkonzepte verwendet werden, die identische Bezeichnungen wie Rollenkonzepte in CTA und VSD tragen (siehe Abb. 44), mit denen jedoch unterschiedliche Interpretationen verbunden.



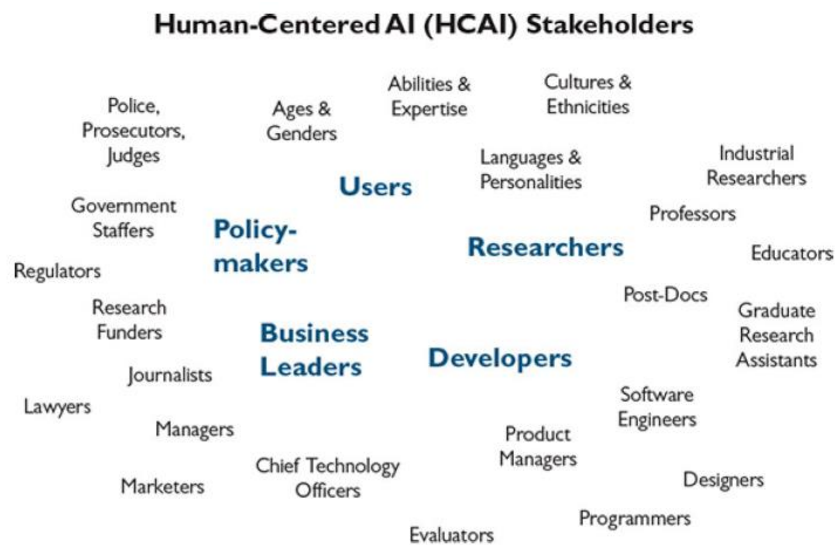


Abbildung 44: Verwendete Rollenkonzepte in HCAI nach Shneiderman (2022, S. 30)

So stehen auch in HCAI Rollenkonzepte im Vordergrund, die menschliches Leben und Handeln in den Lebensbereichen der Gestaltung (Developers) und Nutzung (Users) von Technik verorten, sowie in den Lebensbereichen der Politik (Policymakers) und der Wissenschaft (Researchers). Da in HCAI im Vergleich mit CTA und VSD deutlich tiefere Abstraktionsebenen Anwendung finden, werden bei den dazugehörigen Rollenkonzepten Unterschiede hinsichtlich der mit ihnen verbundenen Interpretationen und Zwecke erkennbar. So zeigt beispielsweise das Attribut „Cultures & Ethnicities“, dass in HCAI bereits Anstrengungen unternommen werden, dem aktuellen Diskurs zu den Implikationen des Nutzerrollenkonzepts Rechnung zu tragen (siehe Kap. 5.6.3). Dies erfolgt in der HCAI-Community aufgrund ihrer im Vergleich mit CTA und VSD jüngeren Historie, ihren Verbindungen zu den Ursprüngen des Diskurses in der HCI-Community, sowie der Bedeutung der Ergebnisse dieses Diskurses für die KI-Gestaltung.

Beim Blick auf die in HCAI verwendeten Rollenkonzepte wird außerdem deutlich, dass eine detailliertere Unterscheidung der Gestalter-Rollenkonzepte vorgenommen wird. So wird bezüglich der Entscheidungsfindung in Gestaltungsprozessen die technische Gestaltung von der strategischen Planung bzw. Entscheidungsfindung durch Rollenkonzepte der „Business Leaders“ oder sinngemäß ähnlicher Bezeichnungen unterschieden, die eine differenzierte Untersuchung und Beschreibung entsprechender Aufgabfelder und Verantwortungsbereiche ermöglicht. Diese erfolgt in CTA und VSD kaum (siehe Kap. 5.6.4). Entsprechend ist in HCAI bezüglich der Unterscheidung und Verzahnung der beiden Verantwortungsbereiche und deren Bedeutung in und für die Gestaltungspraxis mehr Wissen vorhanden. Das Resultat dieses Wissensvorsprungs zeigt sich auch in der Arbeit mit Rollenkonzepten wie den Chief Executive Officers, den Product Managers oder den Security Engineers, die besonders in ökonomischen Kontexten verorteten Gestaltungsprozessen häufig von zentraler Bedeutung sind, und deshalb in HCAI entsprechende Berücksichtigung erfahren. Wie bei anderen Gestalter-Rollenkonzepten fällt auch bei

diesen Rollenkonzepten auf, dass sich ihre Bezeichnung und Charakterisierung stark an in der Realität etablierten Berufsbildern orientiert. Die Untersuchung der Überschneidungen und Unterschiede technischer und organisatorischer Verantwortung auf Detailebene hat sich im Ursprungsfeld HCI zudem als wertvoll erwiesen, da eine Differenzierung zwischen ihnen die Identifikation von Anforderungen erleichtert sowie Ursachen für Prioritäts- oder Zielkonflikte offenlegt. Bezüglich der Detailgrade von Gestalter- und Managerrollenkonzepten in HCAI können CTA und VSD daher kaum Beiträge leisten, jedoch können im Kontext der Auseinandersetzung mit Zielkonflikten die Strategien zur Identifikation und zur Auflösung von Wertekonflikten in VSD wertvolle Perspektiven und Einblicke eröffnen. Denn während sich die Analyse von Prioritäts- und Zielkonflikten in HCAI häufig auf die effiziente Verteilung und Verwendung von Ressourcen bezieht, gehen Wertkonfliktanalysen tiefer und legen Aspekte und Faktoren frei, die solchen Konflikten (möglicherweise unbewusst) zugrunde liegen. Durch die Auseinandersetzung mit solchen Faktoren können in HCAI ganzheitlichere Entscheidungsstrategien ausgearbeitet werden, die die etablierten Sets kritischer Entscheidungsfaktoren vervollständigen. Sie steigern die Wahrscheinlichkeit, Menschen als Lebewesen mit Stärken und Schwächen zu betrachten und zu begreifen, anstelle die in Gestaltungskontexten häufig angewandten mechanistischen Denkweisen von „menschlichem Kapital“ (Auernhammer, 2020, S. 1318). Wertkonfliktanalysen können sich als bedeutsames Werkzeug in der MZKG erweisen, um die angestrebte Menschorientierung zu realisieren, da sie neben ihrem Nutzen für die „internen“ Abläufe in Gestaltungsprozessen besonders essentiell für die Auseinandersetzung mit den Anforderungen, Bedürfnissen und Auswirkungen für direkt und indirekt betroffene Gruppen und Individuen sind. Sie können ferner dazu beitragen, dass in Gestaltungsprozessen anstelle von Symptomen die tatsächlichen Ursachen menschlicher Probleme adressiert werden, wie beispielsweise Donald Norman es von einer menschenzentrierten Gestaltung fordert (siehe Kap. 2.2.3).

Ergänzend zum vorherigen Kapitel zeigen auch die Detailgrade der Nutzer- und Gestalter-Rollenkonzepte, dass in HCAI im Vergleich zu CTA und VSD beim Blick auf Gestaltungsprozesse mehrheitlich noch niedrigere Abstraktionsebenen angewendet werden und dass sich der Trend in HCAI hin zu höheren Abstraktionsebenen erst langsam durchzusetzen beginnt (siehe Kap. 3.3). Während in CTA und VSD Gestaltungsprozesse von einer hohen Abstraktionsebene betrachtet werden, ist in HCAI nach wie vor die konkrete Gestaltungsebene der Ausgangspunkt sämtlicher Betrachtungen. In VSD und CTA kommt eine gedankliche Vogelperspektive zum Tragen, um dem Anspruch einer möglichst holistischen Erfassung von Gestaltungsprozessen gerecht zu werden. In HCAI ist hingegen mehr eine Froschperspektive auf Gestaltungsprozesse vorhanden, die langsam damit beginnt, die aus HCI übernommene Perspektive sukzessive zu erweitern. Dies erfolgt aufgrund der seit den späten 2010er-Jahren einsetzenden Erkenntnis der weiterreichenden Implikationen der Zielsetzungen in HCAI im Vergleich zu den in der MZG formulierten Zielsetzungen. Aufgrund dieser einsetzenden Erkenntnis entsteht in HCAI derzeit eine öffnende Bewegung im Diskurs, die ein Potential darstellt, die in CTA und VSD vorhandenen Erkenntnisse zu menschenzentrierten Gestaltungsprozessen auf höherem Abstraktionsniveau in den HCAI-Diskurs einzubringen und Orientierung bei der Ausarbeitung eigener Positionen und Strategien zu bieten. Der Vergleich der Gestalter- und Nutzerrol-

lenkonzepte zwischen den drei Gestaltungsphilosophien zeigt auch, dass das Einnehmen von höheren Abstraktionsebenen typischerweise den Preis einfordert, Detailgrade aufzugeben. Ein entscheidender Unterschied von HCAI zu VSD und CTA ist jedoch, dass sich das Feld ausschließlich auf die KI-Technologie fokussiert und nicht den weitreichenden Anwendungsanspruch von CTA und VSD auf sämtliche Technologien erhebt. Insofern sind die Betrachtungsgrenzen in HCAI grundsätzlich weniger umfangreich und erfordern eine im Vergleich zu CTA und VSD geringere Erhöhung der gedanklichen Abstraktionsebenen. Dennoch ist diese Erhöhung erforderlich, insbesondere um die in HCAI häufig angenommene Linearität von Abhängigkeiten und Auswirkungen des Verhältnisses von Gestaltung und Nutzung durch die Anerkennung komplexerer Wirkungsketten in HCAI zu erweitern. Um entsprechende Denkweisen zu fördern und in HCAI konstitutiv verankern zu können, können beispielsweise die in CTA verwendeten Analyseebenen der Macro-, Meso- und Microebenen und der ihnen entspringenden Rollenkonzepte aufgegriffen werden, um in HCAI komplexere Abhängigkeitszusammenhänge zu explorieren.

### 5.7.3 Die Bedeutung der Unbeteiligten und Betroffenen

Die ganzheitlichere Betrachtung von Gestaltungsprozessen in VSD und CTA wird auch an den Rollenkonzepten der Unbeteiligten und Betroffenen deutlich. Im Gegensatz zu vielen auf die konkrete Gestaltung fokussierten Ansätzen, werden in VSD und CTA die indirect Stakeholders und die Impactees ins Zentrum der Aufmerksamkeit gerückt. Es wird als eine zentrale Aufgabe angesehen, diese Rollenkonzepte als bedeutsame Zielgruppen zu etablieren, und ihren Eigenschaften und Argumenten in Gestaltungsprozessen eine hohe Aufmerksamkeit zu widmen. Die Auseinandersetzung mit Betroffenen und Unbeteiligten erfährt auch in der KI-Gestaltung zunehmend an Bedeutung, da KI-Systeme immer häufiger Menschenleben auf indirekte Weise beeinflussen, und ihr Einsatz von gesellschaftlichen Herausforderungen und Folgewirkungen hohen Ausmaßes begleitet wird. Somit stellen die Unbeteiligten und Betroffenen auch in HCAI eine besonders bedeutsame Zielgruppe dar, der vor dem Hintergrund des Anspruchs einer menschenzentrierten Gestaltung eine besonders hohe Aufmerksamkeit und Berücksichtigung geschenkt werden sollte. Auch dies erfordert in HCAI eine Erweiterung der bislang angewandten Betrachtungsmaßstäbe auf Gestaltungsprozesse, insbesondere eine Erweiterung der bislang verwendeten Rollenkonzepte. Denkbar ist in diesem Zusammenhang, dass die Unbeteiligten und Betroffenen künftig als eine eigene Rollenkonzeptkategorie in HCAI etabliert werden, oder alternativ als eine ständig vorhandene Subkategorie der Nutzer in HCAI betrachtet werden. Die Antizipation betroffener Gruppen und Individuen ist für HCAI auch insofern förderlich, da sie häufig eine Denkweise **„outside the box“** erfordert. Um sie zu identifizieren müssen alternative Gestaltungs- und Nutzungsszenarien in Erwägung gezogen werden, und sie erfordern eine aufmerksame und gründliche Auseinandersetzung mit dem Gestaltungskontext. Mit ihrer Identifikation ist ein bewusstes Überschreiten gedanklicher Grenzen verbunden, das tendenziell aufwandsreich ist, jedoch den Denkstrukturen vielen als Manager oder Gestalter tätigen Personen in HCAI nicht allzu fernliegt. So ist ein Attribut vieler Gestalter- und Managerrollen ein grundsätzlich antizipatives Denken, das sie benötigen, um beispielsweise mögliche Gefahren des

Einsatzes eines von ihnen gestalteten Systems zu erkennen oder um Schutzmaßnahmen zu realisieren, die böswillige Einsätze ihres Systems verhindern. Das Eingehen der Anstrengungen für die gedanklichen Grenzüberschreitungen, die die Arbeit mit indirect Stakeholders und Impactees erfordern, sind in HCAI somit im Ansatz schon vorhanden und können künftig weiter ausgebaut und verfeinert werden.

Im Aufgreifen der Rollenkonzepte der Unbeteiligten und Betroffenen liegt ein weiterer Mehrwert für die HCAI-Community, da die Rollenkonzepte zusätzlich zur Identifikation zentraler Zielgruppen und ihrer Werte darüber hinaus dazu beitragen, für den gesamten Gestaltungsprozess relevante Chancen und Risiken zu identifizieren. Einmal identifizierte negativ betroffene Gruppen und Individuen ermöglichen es, Präventionsstrategien zur Vermeidung negativer Folgewirkungen zu entwerfen. Erfolgt eine Identifikation von Betroffenen bereits früh im Prozess, kann dem Tätigen von Fehlinvestitionen und weiteren Zuteilungen von Ressourcen vorgebeugt werden, die rückwirkend als nicht zielführend bewertet werden müssten. Dieser Mehrwert bezieht sich auf die mehr ökonomisch geprägte Betrachtungsweise von Gestaltungsprozessen, die in HCAI im Vergleich zu CTA und VSD stärker ausgeprägt ist. Was die Auseinandersetzung mit Unbeteiligten und Betroffenen unter diesem Aspekt zusätzlich lohnenswert macht, ist die Unterstützung der Rollenkonzepte zur Vermeidung bzw. zum reflektierten Umgang mit der sogenannten sunk cost fallacy (D. Friedman et al., 2007; Haita-Falah, 2017; Strough et al., 2008). Die sunk cost fallacy beschreibt das Phänomen der emotionalen Schwierigkeit von Menschen, in Entscheidungsprozessen zu entscheiden, ob ein Projekt fortgeführt oder abgebrochen werden soll, wenn bereits Ressourcen in ein Projekt investiert wurden, diese jedoch nicht die gewünschten Ergebnisse hervorgebracht haben. Der Trugschluss (fallacy) bezieht sich dabei auf die Argumentation, dass allein aufgrund der Tatsache bereits getätigter Investitionen das mit einem Projekt verfolgte Ziel weiterverfolgt werden sollte, auch wenn zum gegenwärtigen Zeitpunkt bereits ersichtlich ist, dass das Ziel nicht (mehr) auf die ursprünglich angedachte Weise erreicht werden kann. Dieses Phänomen und der unreflektierte Umgang damit ist in vielen TG-Kontexten die Ursache für das Scheitern von Projekten und bezieht sich häufig auf Faktoren wie beispielsweise gestiegene Energiekosten, die veränderte Verfügbarkeit von Rohstoffen oder von qualifiziertem Personal. Auch in der KI-Gestaltung ist die Gefahr einer „Versenkung“ von Ressourcen von hoher Relevanz, da die zur Gestaltung von KI-Systemen aufzubringenden Ressourcen teilweise beträchtliche Größenordnungen erreichen, die ein weit vorausschauendes Ressourcenmanagement erfordern. Die Übernahme der Rollenkonzepte der Impactees und indirect Stakeholders kann dies unterstützen, da sie helfen, frühzeitig negative Auswirkungen von KI-Systemen aufzuzeigen, die durch entsprechende Entscheidungsfindungen und Umlenkungen von Ressourcen vermieden werden können. Die kontinuierliche Auseinandersetzung mit Unbeteiligten und Betroffenen kann somit in HCAI das Risiko senken, Fehlinvestitionen zu tätigen und sich mit der sunk cost fallacy auseinandersetzen zu müssen (oder ihr gar zu erliegen). Die Rollenkonzepte besitzen für die HCAI-Community einen Synergieeffekt, da sie zusätzlich zu den von ihr primär adressierten sozialen und ethischen Aspekten sekundär auch Managementaspekte in Gestaltungsprozessen berücksichtigen helfen. Da in HCAI das Denken in Managementperspektiven bereits etabliert ist, ist denkbar, dass der mit den Rollenkonzepten verbundene Preis einer geringen Aufwandssteigerung zu Beginn

von Gestaltungsprozessen zugunsten der Vermeidung hoher Kosten zu fortgeschrittenen Zeitpunkten, auf Einsicht stoßen. Nicht zuletzt auch deshalb, weil ähnliche Gedankengänge in seinem Ursprungsfeld der HCI bereits gutgeheißen wurden, als in den 1980er- und 1990er-Jahren im Kontext der Softwareentwicklung sich der Paradigmenwechsel von linearen zu zyklischen Prozessmodellen vollzog (siehe Kap. 2.1).

#### 5.7.4 Rollenkonzepte des Data Managements

Aufgrund der Fokussierung der KI-Technologie stehen in HCAI einige Rollenkonzepte im Vordergrund, die in CTA und VSD keine Erwähnung finden. Zu diesen zählen primär Rollenkonzepte aus dem Fachbereich des Data Managements, etwa wie dem Data Cleansing, dem Data Mining oder der Data Science (Ozkaya, 2020). Rollenkonzepte aus diesem Fachbereich beschreiben Gruppen und Individuen, die sich durch verschiedene Tätigkeiten mit der Erhebung und Aufbereitung digitaler Daten auseinandersetzen. Beispielsweise bereinigen und sichern sie Datenmaterial und schaffen Strukturen, mit deren Hilfe Daten durch KI-Systeme abgerufen und verarbeitet werden können. Derartige Rollenkonzepte sind in HCAI von zentraler Bedeutung, da im Kontext der Gestaltung von KI-Systemen Gestaltungserfolge zunehmend an die Verfügbarkeit, Menge und Qualität von Daten geknüpft sind. Aufgrund der Bedeutung von Daten wird deshalb häufig auch von „datengetriebenen“ (data-driven) Gestaltungs- und Entwicklungsverfahren in der KI-Gestaltung gesprochen (Ntoutsis et al., 2020). Durch die Bedeutsamkeit von Daten sehen sich auch in HCAI aktive Entscheidungsträger und Gestalter mit neuen Herausforderungen konfrontiert, denn die zunehmend datengetriebenen Gestaltungsverfahren unterscheiden sich von den mehr problemorientierten (problem-driven) Vorgehensweisen vieler agiler Softwareentwicklungsverfahren, auf die das Feld HCI und der Ansatz MZG traditionell stark Bezug nehmen. Der Rückgriff auf das im Kontext von Softwareentwicklung etablierte Wissen kann deshalb oftmals nur bedingt Aufschluss darüber geben, wie die neuen durch die KI-Technologie hervorgerufenen Herausforderungen bewältigt werden können. Ein konkretes Beispiel für eine solche Herausforderung sind die in großen Datenmengen vorhandenen Verzerrungen (Biases), die es für Data Manager zu identifizieren und aufzulösen gilt. Ihre Aufgabe ist es, Möglichkeiten zur Abwendung der Risiken für Diskriminierung zu entwickeln, die aus der Verwendung solch unbereinigter Daten resultieren können. So berühren auch die Tätigkeiten von Rollen im Data Management häufig ethische Aspekte und Fragestellungen, die es für die in HCAI verfolgten Ziele zu adressieren gilt. Erschwerend kommt bei dem genannten Beispiel hinzu, dass die Auseinandersetzung mit Biases oftmals über den gesamten Verlauf von Gestaltungsprozessen, stellenweise auch darüber hinaus erfolgen muss, da die Datenmengen stetig anwachsen und eine entsprechende Pflege erfordern. Zur angemessenen Adressierung solcher Aspekte können die Gestaltungsprinzipien der MZG in HCAI jedoch kaum Hilfestellung leisten, da sie weder konkrete Aspekte zum Datenmanagement noch zum Umgang mit ethischen Fragestellungen ausreichend umfangreich thematisieren. In HCAI müssen deshalb neue Lösungen entwickelt, möglicherweise auch neue Gestaltungsprinzipien formuliert werden, die die in der MZG angebrachten Betrachtungsgrenzen überschreiten. Dass in der CTA- und VSD-Literatur Rollenkonzepte zum Data Management keine Erwähnung finden, erklärt sich

schlicht durch den Blick auf die jüngere Geschichte der KI-Gestaltung – Solche Rollen und die Bedarfe für die in ihnen beschriebenen Tätigkeiten entstehen durch den jüngsten technischen Fortschritt der praktischen KI-Gestaltung, die auf Detailebene in der VSD- und CTA-Literatur bislang noch keine Widmung erfahren hat.

### 5.7.5 Strukturelle Unterschiede zwischen der Gestaltung von Software- und KI-Systemen

Zusätzlich zur gestiegenen Bedeutung von Daten zeichnet sich die Gestaltung von KI-Systemen durch weitere strukturelle Unterschiede im Vergleich zur Softwareentwicklung aus, die Implikationen für Denkweisen in HCAI besitzen. Beispiele hierfür sind der stärkere Fokus auf die Entwicklung von Modellen und Algorithmen sowie die für die Gestaltung der Systeme notwendige Expertise in Fachbereichen wie der Statistik oder des Machine Learning. Software wird in der Regel gestaltet, um vordefinierte Aufgaben auszuführen, während in der KI-Gestaltung Algorithmen und Verfahren entwickelt werden, um menschenähnliches Denken zu simulieren und um stärker auf kontextspezifische Anforderungen eingehen zu können. KI-Systeme stellen hochdynamische Systeme dar, die ihre interne Struktur und ihre Funktionalitäten häufig stark wandeln. Im Vergleich dazu ist es bei Softwarelösungen ohne oder mit nur geringen KI-Anteilen üblich, dass sie nach ihrer Einführung nur punktuell und zu definierten Zeitpunkten Updates erhalten. Solche strukturellen Unterschiede führen auch dazu, dass sich im Einsatz von KI-Systemen neue Mensch-Technik-Verhältnisse ausbilden. Dadurch sehen sich in HCAI viele Gestalter und Entscheidungsträger mit der Herausforderung konfrontiert, ihr Verständnis der Beziehungen zwischen Mensch und Technik, im Besonderen die Bedeutung der Mensch-Maschine-Interaktion zu reflektieren und weiterzuentwickeln. So werden viele der Zustandsänderungen von KI-Systemen häufig von Menschen hervorgerufen, die mit den Systemen interagieren, die jedoch nicht an der Gestaltung der Systeme beteiligt waren oder als potenzielle Nutzer identifiziert worden waren – ein weiterer Gesichtspunkt, der für den Mehrwert der Arbeit mit den Rollenkonzepten der Unbeteiligten und Betroffenen in HCAI spricht. Aus Interaktionssituationen werden neue Daten generiert, ausgewertet und in neuen Versionen von KI-Systemen verwendet, um die Systeme besser an die Anforderungen und Arbeitsabläufe von Menschen anzupassen. Solche „Interaktionsschleifen“ können in der Nutzung von KI-Systemen beliebig viele Zyklen durchlaufen. Dadurch verliert die in der MZG häufig angewandte, mehr abgegrenzte Perspektive auf Gestaltung und Nutzung an Tragfähigkeit und erfordert in HCAI eine Anpassung hin zu Perspektiven, die weitreichende Zeithorizonte, Wirkungsketten und Interaktionsmodelle in Betracht ziehen und berücksichtigen (Ahmad et al., 2021; Heyn et al., 2021). Auch zeigt dieser Umstand, weshalb eine Revision der mit vielen Nutzerrollenkonzepten verbundenen Vorstellungen eines Endverbrauchers im Kontext der KI-Gestaltung angebracht ist, da **die „Enden“ von Wertschöpfungsketten** deutlich weniger klar und final definiert werden können, wie es in den Methoden der Softwareentwicklung üblich ist. Im Hinblick auf die Perspektive des Verhältnisses von Gestaltern und Nutzern in CTA und VSD hat die vergleichende Analyse gezeigt, dass in beiden Ansätzen auch eine im Ansatz traditionelle und rigide Auffassung des Verhältnisses zum Tragen kommt. Aufgrund der in beiden Ansätzen vor-

handenen Offenheit und Flexibilität der dazugehörigen Rollenkonzepte ist ihre Auffassung jedoch tendenziell weniger eng bemessen. Ein Blick auf die Interpretationen der Gestalter- und Nutzerrollenkonzepte in CTA und VSD kann in HCAI daher dabei helfen, die Ausbildung eines neuen (Selbst-)Verständnisses über die Verhältnisse von Mensch und KI-Systemen sowie die Entwicklung angepasster Perspektiven auf die Interaktion zu entwickeln und zu etablieren. Durch die Bedeutung von Interaktion in VSD sind die dazugehörigen VSD-Rollenkonzepte hierfür prädestiniert.

Auch die Rollenkonzepte der Unbeteiligten und Betroffenen erfahren unter dem Gesichtspunkt der strukturellen Unterschiede von Software- und KI-Gestaltung in HCAI an Relevanz. Im Vergleich zu Software ohne KI-Anteile werden KI-Systeme häufiger explorativ gestaltet, was bedeutet, dass die Gestaltung beginnt, ohne bereits im Vorfeld eines Gestaltungsprozesses konkrete Zielgruppen identifiziert oder Verwendungszwecke antizipiert zu haben. Hierdurch entsteht in HCAI ein potenzieller Konflikt mit den Gestaltungsprinzipien der MZG, denn sie sehen die Integration von Nutzern und das Durchführen von Requirements Engineering Methoden als wesentliches Element menschenzentrierter Gestaltung vor. In mehr explorativen KIGP und häufig auch in frühen Gestaltungsphasen nicht-explorativer KI-Systeme fehlen jedoch häufig die dafür notwendigen grundlegenden Informationen zu möglichen Nutzergruppen. Deshalb gewinnen sämtliche Tätigkeiten zur frühzeitigen Identifikation möglicher Nutzer und ihrer Verhaltensweisen, Werte und Bedürfnisse (User Research) in HCAI stark an Bedeutung (Baxter et al., 2015). Hinzu kommt, dass es sich bei den Nutzergruppen im HCAI-Kontext um weit weniger klar abgrenzbare Gruppen handelt, da sich beispielsweise selbstlernende KI-Systeme an verschiedene Kontexte und Bedürfnisse von Gruppen und Individuen anpassen können. Die Identifizierung und Auseinandersetzung mit möglichen direkten und indirekten Stakeholdern eines Systems muss daher in menschenzentrierten KI-Gestaltungsprozessen zu verschiedenen Zeitpunkten und deutlich kontinuierlicher durchgeführt werden, um dem Anspruch menschenzentrierter Gestaltung in HCAI gerecht werden zu können. Eine Möglichkeit zur Bewältigung dieser Herausforderung kann es sein, hierfür die Erfahrungswerte und Strategien aus CTA und VSD aufzugreifen und für die spezifischen Anforderungen in HCAI zu modifizieren und weiterzuentwickeln. Beispielsweise können hierfür die Methoden der Technologiefeldanalyse aus CTA oder die Methoden der conceptual investigations aus VSD zur Ermittlung von direkt und indirekt betroffenen Stakeholdern aufgegriffen werden.

Regelmäßig durchgeführt, können sie dazu beitragen, dass in HCAI Unbeteiligte und Betroffene während der gesamten Dauer von Gestaltungsprozessen identifiziert und im fokussiert werden können. Mittel- bis langfristig können so auch Anpassungen und Ergänzungen der Gestaltungsprinzipien der MZG für HCAI erfolgen.

### 5.7.6 Die Verwendung von Rollenkonzeptpaaren

Die Analyse der verwendeten Rollenkonzepte in CTA und VSD hat auch gezeigt, dass in beiden Ansätzen häufig mit Rollenkonzeptpaaren gearbeitet wird. Neben der mehr übergreifenden Paarung aus Gestaltung und Nutzung finden sich in CTA und VSD die promi-



nenten Rollenkonzeptpaare der Impactors und Impactees, der Enactors und comparative Selectors, oder der direct und indirect Stakeholders. Übergeordnet erfüllen diese Rollenkonzeptpaare in beiden Ansätzen den Zweck, die Komplexität sozialer Strukturen zu reduzieren. Auf konkreter Ebene erfolgt mit ihnen häufig eine Gegenüberstellung von verschiedenen Positionen in spezifischen Aushandlungsfragen im Gestaltungsgeschehen. Rollenkonzeptpaare wie beispielsweise die Technikenthusiasten und Technikgegner zeigen zwei Pole auf, um gegensätzliche Interessenslagen darzustellen und zu verdeutlichen. Stellenweise werden mit ihrer Hilfe auch manche Umstände bewusst überzeichnet (Metapher von Gewinnern und Verlierern), um die Chancen des Berücksichtigens und die Risiken des Ignorierens von Fragen zu sozialen und ethischen Aspekten in Gestaltungsprozessen hervorzuheben. Die Rollenkonzeptpaare erlauben es, Grenzlinien zwischen gedanklichen Konstrukten zu ziehen, um die Notwendigkeit oder Angemessenheit der in CTA und VSD vorgeschlagenen Methoden darzulegen. Dabei erschließt sich die volle Bedeutung eines der beiden Rollenkonzepte häufig nur in der zusammenhängenden Betrachtung des Gegenstück-Rollenkonzepts. Da das Bewusstsein für die Bedeutung ethischer und sozialer Aspekte im Gestaltungsgeschehen in HCAI bereits etabliert ist, kann die Adaption von Rollenkonzeptpaaren mit ihrer Funktion der Gegenüberstellung hilfreich sein, um die teilweisen komplexen Geflechte der an Gestaltungsprozessen beteiligten Interessensgruppen zu entwirren und ein besseres Verständnis für die mehr verborgenen Motivationen von Stakeholdern zu erlangen. Allerdings sollte dabei beachtet werden, dass insbesondere bei der Auseinandersetzung mit Wertefragen kein voreiliges und kategorisches Schwarz-Weiß-Denken angewendet wird, das manche der Rollenkonzeptpaare bei einem zu ungenauen Blick suggerieren können. Die Rollenkonzeptpaare erlauben es, den verschiedenen Spektren und Abstufungen („Grauschattierungen“) gedanklich wie faktisch besondere Behutsamkeit entgegenzubringen, die besonders in tendenziell heiklen Wertefragen von Bedeutung ist (siehe Kap. 5.6.7).

### 5.7.7 Mehrwert für die Untersuchung ethischer Aspekte

Im Vergleich zu den traditionellen Fragestellungen im Feld HCI stellt die Notwendigkeit, sich deutlich intensiver mit ethischen Implikationen gestalterischen Handelns auseinanderzusetzen, ein zentrales Novum für die HCAI-Community dar. KI-Systeme berühren aufgrund ihrer technisch höheren Komplexität und ihren vielfältigen Einsatzfeldern in deutlich kritischerem Umfang Wertefragen, die in HCAI vor dem Hintergrund des Anspruchs einer menschenzentrierten Gestaltung adressiert werden müssen. In seiner Ursprungsdisziplin der HCI werden bereits punktuell ethische Fragestellungen andiskutiert, jedoch erfolgt der Zugang zu diesen häufig über psychologische Perspektiven, die sich der Thematik mit Fragestellungen und Erkenntnisinteressen zu Emotionen annähern. Zwischen Emotionen und Werten kann eine Abhängigkeit bestehen (Breitsameter, 2015; Steinfath, 2001), die Betrachtung dieser erfährt jedoch für viele Fragestellungen in der HCI-Community aufgrund höher priorisierter gestalterischer Fragen üblicherweise nur geringe Aufmerksamkeit. Ferner ist in der HCI-Community vereinzelt auch ausgeprägte philosophische Expertise vertreten, die angesichts der Erkenntnis der Bedeutung ethischer Fragestellungen in den vergangenen 20 Jahren zusehends gestärkt und abgerufen wurde. Nicht zuletzt der



Ansatz VSD ist ein Beispiel hierfür. Dennoch besteht in HCAI nach wie vor ein hoher Bedarf an Expertise und Erfahrungswerten, die Orientierung zum zielführenden Umgang mit ethischen Fragestellungen in Gestaltungskontexten bieten können. Zur Deckung dieses Bedarfs kann sich eine Auseinandersetzung mit sämtlichen Rollenkonzepten in CTA und VSD lohnen, besonders mit den Argumentationslinien, die ihrer Entwicklung und Verwendung zugrunde liegen. Tendenziell halten für HCAI die VSD-Rollenkonzepte mehr Hilfestellungen wie jene in CTA bereit, da in CTA die Förderung von Reflexionsarbeit und weniger die tatsächliche Gestaltung im Vordergrund steht. Insbesondere die Auseinandersetzung mit dem Rollenkonzept der CTA agents kann in HCAI ausgespart werden, da sie dem eindeutig zum Ausdruck gebrachten Anspruch von direkter und aktiver Gestaltung in HCAI entgegenstehen (siehe Kap. 5.6.5). In VSD wird hingegen die Auseinandersetzung mit ethischen Fragestellungen in Verbindung mit klaren Gestaltungszielen angestrebt (progress not perfection). Die darin verankerte Erkenntnis, dass in der Auseinandersetzung mit ethischen Fragen im Zweifel auch Kompromisse zugunsten des Gestaltungsfortschritts eingegangen werden müssen, verspricht besonders wertvolle Impulse für den MZKG-Diskurs in HCAI. Tatsächlich erfolgen aus dieser Motivation heraus bereits erste Versuche, Gedanken und Konzepte von VSD mit den Zielen in HCAI zu verknüpfen (Aldewereld & Mioch, 2021; Sadek et al., 2023; Schoenherr et al., 2023). Ein weiterer konkreter Mehrwert in der Auseinandersetzung mit ethischen Fragestellungen sind sicherlich auch die Rollenkonzepte der Betroffenen aus CTA und VSD, da durch ihre Verwendung die potenziellen und tatsächlichen negativ von KI-Systemen betroffenen Gruppen und Individuen identifiziert werden können, denen im Kontext der MZKG sehr hohe Aufmerksamkeit gewidmet werden muss. Diese Rollenkonzepte können dazu beitragen, frühzeitig mögliche Konflikte und Schäden zu identifizieren, die durch die Ausbildung vorbeugender Maßnahmen abgewendet werden können. Auch halten die in den Rollenkonzepten der informierten Einflussnehmer thematisierten Machtgefüge aus CTA aufschlussreiche Perspektiven für HCAI bereit, da die Auseinandersetzung mit den (nicht direkt erkennbaren) Machtstrukturen in KI-Gestaltungssettings und somit mehr politischen Aspekten tendenziell mehr Raum einnimmt, wie in Prozessen der Softwareentwicklung.

Die Frage, inwieweit in HCAI ethische Expertise aufgebaut und verortet werden soll und muss, kann in 2024 noch nicht final beantwortet werden. So hat sich parallel zur Entwicklung des Feldes HCAI auch das Feld der Ethical AI herausgebildet (Mittelstadt, 2019; D. S. Rubenstein, 2021; Siau & Wang, 2020), in dem bereits intensiv Fragestellungen zu ethischen Implikationen von KI und auch KI-Gestaltung erörtert werden (siehe Kap. 3.2.2). So wird in 2024 beispielsweise in verschiedenen Diskursen noch ausgehandelt, inwieweit eine menschenzentriert gestaltete KI notwendigerweise auch eine nach ethischen Maßstäben gestaltete KI darstellen muss (und umgekehrt). Es steht im multidisziplinären Diskurs der Communities die Frage nach den Grenzen und Überschneidungen der Verantwortungsbereiche im Raum, die bislang noch nicht eindeutig abgesteckt werden konnten und möglicherweise auch nicht zwingend voneinander getrennt werden müssen. Insofern ist es für den derzeitigen Diskurs in HCAI von Vorteil, dass sowohl in HCAI wie auch in der Ethical AI-Community sich mit ethischen Fragestellungen zu KI auseinandergesetzt wird, da beide Felder eindeutig traditionell stark ausgeprägte Expertisen besitzen, die

beide zur zielführenden Bewältigung der Fragestellungen der MZKG beitragen. In der Positionierungsfrage, wie tiefgehend in HCAI ethische Fragestellungen adressiert werden können und müssen, kann eine Auseinandersetzung mit den Rollenkonzepten in CTA und VSD ebenso zielführend sein. Eine Übernahme und Verwendung der Rollenkonzepte von beispielsweise den Impactors und Impactees kann zeigen, ob die Rollenkonzepte den Anforderungen zur Adressierung ethischer Fragestellungen in HCAI gerecht werden können. Dies wird sich auch in der Zukunft als bedeutsame Aufgabe in HCAI erweisen, da derzeit auch diskutiert wird, „how human-centered also means humanity centered“ (F. Dignum & Dignum, 2020, S. 1). So spricht beispielsweise auch Donald Norman in seiner Auslegung der MZG von 2023 nicht mehr nur von menschenzentrierter, sondern von menscheitszentrierter (humanity centered) Gestaltung (Norman, 2023).

### 5.7.8 Verwendung verschiedener Abstraktionsebenen

Die Diskussion war bis hierher stark von dem Thema geprägt, dass in HCAI, ähnlich zu CTA und VSD, höhere gedankliche Abstraktionsebenen benötigt werden, um den dort hervortretenden komplexeren Fragestellungen und neuen Ausprägungen von Mensch-Technik-Verhältnissen begegnen zu können. Dieser Umstand soll jedoch nicht darüber hinwegtäuschen, dass trotz der Notwendigkeit für die Entwicklung und Adaption von (neuen) Rollenkonzepten, in HCAI bereits etablierte Stärken vernachlässigt oder gar aufgegeben werden sollten. Dies betrifft insbesondere die aus dem Feld HCI übernommenen Abstraktionsebenen zu menschlichem Leben und Handeln, durch deren Anwendung in der Vergangenheit bereits Gestaltungserfolge erzielt werden konnten. Ein wesentlicher Erfolgsfaktor von HCI als Feld und auch von der MZG als Gestaltungsansatz liegt darin, dass sie Entscheidungsträgern die Möglichkeit bieten, sich auf sehr tiefe Abstraktionsebenen zu begeben und menschliches Handeln sehr konkret zu analysieren. Dazu wird der Mensch häufig in seinen Rollen in kleinen sozialen Gruppen, nicht selten auch als Individuum betrachtet. Die Untersuchung von Einstellungen und Werten von Individuen auf solch tiefen Abstraktionsebenen ermöglicht es, wertvolles Wissen und Schlüsselaspekte zu Wertefragen und Wertekonflikten freizulegen. Auch ermöglichen sie es, im Sinne der MZG die möglicherweise besonderen Anforderungen und Bedingungen relevanter Gruppen und Individuen berücksichtigen zu können. Diese Aspekte haben in HCAI nach wie vor Bestand und gewinnen insbesondere unter dem Gesichtspunkt ihrer Beitragsfähigkeit zu ethischen Fragestellungen vielmehr noch an Gewicht. Entsprechend besteht für HCAI eine Chance darin, künftig Methoden zu entwickeln, die ein anlassbezogenes und für die Gestaltung von KI-Systemen zielführendes Wechseln zwischen konkreten und abstrakten Rollenkonzepten, zwischen tiefen und höheren Abstraktionsebenen ermöglicht. Die Rollenkonzepte aus CTA können hierfür kaum Hilfestellung leisten, da tiefe Abstraktionsebenen und besonders das Individuum in der soziologisch geprägten Betrachtungsweise von Gestaltungsprozessen in CTA ausgespart werden. In VSD hingegen finden sich vereinzelt Hinweise darauf, dass die Bedeutung niedriger Abstraktionsebenen bekannt ist und ihr Mehrwert geschätzt wird. So widmen sich einige für VSD bedeutsame Studien beispielsweise körperlich eingeschränkten Gruppen und Individuen (Azenkot et al., 2011; Deibel, 2011; Elkind et al., 1996) – ein Bezugspunkt zum in HCI traditionell bedeutsamen

Themenfeld der Barrierefreiheit (Accessibility), welches auch in HCAI Relevanz erfährt. Ferner werden in VSD auch Erkenntnisse aus pädagogischen Feldern aufgegriffen, beispielsweise um Wertekonflikte bei der Einführung von IT-Systemen in Organisationen zu untersuchen. Somit werden in VSD verschiedene ergonomische, psychologische und pädagogische Erkenntnisse aufgegriffen, die die Betrachtung des Menschen als Individuum nahelegen. Ein weiterer Zugang zu tiefen Abstraktionsebenen erfolgt in VSD durch die Referenz der philosophischen Schule des Humanismus, beispielsweise im Zusammenhang mit der Methode der Envisioning Cards (B. Friedman & Hendry, 2012). Die Auseinandersetzung mit humanistischen Gedanken und Ideen findet auch im Diskurs zwischen der HCAI-Community und der Ethical AI-Community statt, insbesondere unter dem Gesichtspunkt, dass sie häufig sowohl die Handlungsfähigkeit von gesellschaftlichen Gruppen wie auch Individuen thematisieren. Insofern kann eine vertiefte Auseinandersetzung mit den humanistischen Gedanken die in VSD aufgegriffen und operationalisiert werden sich in doppelter Hinsicht fruchtbar für HCAI erweisen, da es einerseits die Verwendung verschiedener Abstraktionsebenen fördert und gleichzeitig den MZKG-Diskurs informiert.

## 5.8 Zusammenfassung

Die Analyse der verwendeten Rollenkonzepte in CTA und VSD, ihrer Ursprünge sowie blinden Flecke hat gezeigt, wie vielfältig menschliches Leben und Handeln in den beiden Ansätzen in Kontexten von TG betrachtet und für diese operationalisiert wird. Die dazu eingenommenen Abstraktionsebenen variieren voneinander, jedoch wird in beiden Ansätzen zwischen verschiedenen Ausprägungen von Betroffenen, Gestaltern, Entscheidungsträgern sowie weiteren als relevant erachteten Beteiligten und Involvierten unterschieden. In beiden Ansätzen werden durch die Wahl der Abstraktionsebenen Kompromisse eingegangen, die jedoch überwiegend transparent dargestellt und begründet werden, weshalb beide Ansätze auch in 2024 Potenziale besitzen, die Betrachtungen menschlichen Lebens und Handelns im MZKG-Diskurs zu bereichern. Insbesondere die Gedanken und Perspektiven zur Auseinandersetzung mit den Rollenkonzepten der Unbeteiligten und Betroffenen, sowie der Technikgegner besitzen für den MZKG-Diskurs besondere Relevanz, da sie, auch aufgrund bereits stattfindender gesellschaftlicher Entwicklungen, stärker in die Konzeption und Planung von MZKG-Prozessen Eingang finden müssen. Auch die in beiden Ansätzen vertretene und mehrfach ausdrücklich dargelegte Position, dass bei der Realisierung von mensch- und wertorientierten technischen Systemen Erkenntnisse und Ergebnisse in verschiedenen Färbungen, Abstufungen und Spektren zu erwarten sind, kann aktuell stattfindende Lernprozesse in der HCAI-Community unterstützen und dazu beitragen, Abstand von der MZKG unzuträglichen binären Mindsets und Gestaltungshaltungen zu wahren. Trotz der zur Realisierung der MZKG-Ziele notwendigen Öffnung auf konzeptioneller Ebene gilt es für die HCAI-Community, ihre ausgewiesenen Stärken nicht zu vernachlässigen. Dies bedeutet zusätzlich zur Adaption eines Denkens in weitreichenden Zeitfenstern und Ursache-Wirkungs-Ketten von Mensch-Technik-Beziehungen die Möglichkeit zu wahren, den Menschen als das zu sehen was er ist – ein Lebewesen mit individuellen Interpretationen, Bedürfnissen, Stärken und Schwächen.

Die vergleichende Analyse hat darüber hinaus gezeigt, dass Rollenkonzepte in beiden Ansätzen auf sprachlicher Ebene dazu verwendet werden, Gedanken zu klassifizieren und zu strukturieren, weshalb die Rollenkonzepte einen wichtigen Bestandteil der CTA- und VSD-Taxonomie darstellen. Ein Stilmittel, das in beiden Ansätzen verwendet wird, sind die Rollenkonzeptpaare, die durch ihren dualistischen Charakter kontrastreich menschliches Leben und Handeln zu erschließen helfen, auch wenn hierfür die Realitätszustände zu Erkenntniszwecken stellenweise bewusst überspitzt oder stark reduziert werden (Modellcharakter). Hinzu kommt, dass insbesondere zu den Rollenkonzepten in CTA auf sprachlicher Ebene nicht alle Mehrdeutigkeiten ausgeräumt werden. Sollten sich Mitglieder der HCAI-Community dazu entschließen, Erkenntnisse aus CTA und VSD in den Diskurs zur MZKG aufzugreifen, empfehle ich deshalb vorher die genauen Bedeutungen und Auslegungen der Rollenkonzepte zu studieren, um etwaigen Missverständnissen vorzubeugen.

Zu den konzeptionellen Fragestellungen der MZKG sind die CTA- und VSD-Rollenkonzepte anschlussfähig, da sie Gestaltungsprozesse und die darin aktiven sowie die davon betroffenen Gruppen und Individuen ganzheitlich erfassen. Die verschiedenen Ausprägungen der Rollenkonzepte können in der MZKG-Debatte dazu beitragen, relevante Stake-

holdergruppen, ihre Werte und Bedürfnisse zu identifizieren und sie angemessen in KIGP zu berücksichtigen. Zur Auffassung von Menschzentriertheit, welche in der MZKG vertreten wird, sind die Rollenkonzepte überwiegend anschlussfähig, stellenweise auch deckungsgleich. Einzige Ausnahme stellt hier das Rollenkonzept der CTA agents dar, da auf die in ihm zum Ausdruck gebrachte, distanzwahrende Mittlerrolle zu wenig belastbare Verantwortung übertragen werden kann und es die für die MZKG notwendige, gestaltungspragmatische Einstellung vermissen lässt. Der MZKG-Diskurs strebt danach, Gestalter und Entscheidungsträger auch in Zeitdrucksituationen entscheidungsfähig zu machen, weshalb sich die HCAI-Community auch intensiv mit der Entwicklung von Gestaltungsmethoden auseinandersetzt. Der Frage, inwieweit die Methoden aus CTA und VSD zu diesem Gesichtspunkt Beiträge leisten können, widme ich mich deshalb im Anschluss.

## 6 Methodik von CTA und VSD

Die gestaltungspragmatische Einstellung im MZKG-Diskurs zeichnet sich besonders dadurch aus, dass die HCAI-Community aktiv Anstrengungen unternimmt, Gestaltungsframeworks und Gestaltungsmethoden zu entwickeln, die einen Transfer der konzeptionell erarbeiteten Erkenntnisse in die Gestaltungspraxis ermöglichen. Neu für die Community ist dabei die Herausforderung, dass zur Realisierung der MZKG-Philosophie soziale und ethische Gestaltungsfaktoren deutlich stärker berücksichtigt werden müssen, was im Ergebnis – im Vergleich zu Modellen wie der MZG – in deutlich umfassendere und deutlich komplexere Gestaltungsframeworks mündet (siehe Kap. 3.3). Zum Umgang mit dieser Komplexitätssteigerung und ihrer Implikationen für die Entwicklung und Anwendung von Gestaltungsmethoden kann es hilfreich sein, Ideen, Strategien und Vorschläge aufzugreifen, die in der Vergangenheit mit vergleichbarem Ansinnen entwickelt wurden. Durch die Bestrebungen, sozial- und wertorientierte Gestaltungsprozesse zu fördern, erfüllen die Methoden aus CTA und VSD dieses Kriterium, weshalb sie sich für eine tiefergehende Untersuchung im Kontext des MZKG-Diskurses eignen. Der Untersuchung der Gestaltungsmethoden von CTA und VSD ist daher dieses Kapitel gewidmet. Zu Beginn werde ich meine Vorgehensweise zur Untersuchung der Gestaltungsmethoden beschreiben (siehe Kap. 6.2). Ich führe eine Expertenevaluation durch, deren Inhalte ich mithilfe einer SWOT-Analyse auswerte. Dies ermöglicht es mir, die von KI-Experten identifizierten Stärken (siehe Kap. 6.3), Schwächen (siehe Kap. 6.4), Chancen (siehe Kap. 6.5) und Risiken (siehe Kap. 6.6) der Methoden strukturiert darzulegen und anhand dieser die Potenziale (Forschungsfrage 4) und die Defizite (Forschungsfrage 5) der Methoden für den HCAI-Diskurs zu ermitteln. Unter Anwendung von im HCAI-Diskurs relevanten Kriterien diskutiere ich anschließend Erkenntnisse der Expertenevaluation (siehe Kap. 6.7). Dies ermöglicht mir zusätzlich, Potenziale zur Weiterentwicklung der Methoden unter gestaltungspragmatischen Gesichtspunkten aufzuzeigen (Forschungsfrage 6).

### 6.1 Einführung

Für die Zielsetzungen der MZKG sind Gestaltungsmethoden von zentraler Bedeutung. In der Gestaltungspraxis unterstützen sie Gestalter und Entscheidungsträger bei der Identifikation bedeutsamer Gestaltungsfaktoren, im reflektierten Treffen von Gestaltungsentscheidungen und bei der Planung und Koordination von Gestaltungsprozessen. Gestaltungsmethoden sind die Werkzeuge, die es ermöglichen, soziale und ethische Kontextfaktoren in zentrale Orientierungsgrößen für Gestaltungsprozesse zu übersetzen und ihren Status im Gestaltungsverlauf konstant im Blick zu behalten. Durch ihren Anleitungs- und Kontrollcharakter tragen die Methoden zum systematischen Gestaltungsvorgehen bei und unterstützen Gestalter und Entscheidungsträger dabei, die positiven Potenziale der KI-Technologie zu heben, die negativen Potenziale zu identifizieren und Strategien zu ihrer Abwendung zu formulieren. An Gestaltungsmethoden ist ein funktionaler, aber auch ein ideeller Anspruch geknüpft, denn sie sollen die theoretisch erarbeitete

ten Standpunkte zur MZKG für die Gestaltungspraxis operationalisieren. Sie stellen das zentrale Bindeglied zwischen Gestaltungstheorie und Gestaltungspraxis dar, da erst durch die Anwendung von Gestaltungsmethoden die in der MZKG angestrebte Menschzentriertheit in KI-Systemen realisiert und evaluiert werden kann. Ein Blick auf den HCAI-Diskurs in 2024 zeigt, dass die Ausarbeitung der HCAI-Gestaltungsphilosophie kontinuierlich Fortschritte verzeichnet, die Auseinandersetzung mit ihrer Operationalisierung jedoch vernachlässigt wird (Xu, Dainoff, et al., 2023). Dies führt dazu, dass in der Gestaltungspraxis ein Mangel an Gestaltungsmethoden vorherrscht, der es erschwert, die in der HCAI-Community entwickelten Ideen zu einer MZKG praktisch zu realisieren. So haben Vertreter der Gestaltungspraxis verschiedentlich um Hilfestellungen zur Umsetzung der konzeptuellen HCAI-Gedanken gebeten, die deutlich stärker die besonderen Anforderungen der praktischen KI-Gestaltung berücksichtigen (Ahmad et al., 2021; Mazarakis et al., 2023; Xu & Gao, 2023). Insofern ist es wenig verwunderlich, dass Unternehmen, die sich der Gestaltung von KI-Systemen widmen, bislang weit hinter den Gestaltungsidealen und Resultatserwartungen der HCAI-Community hinterherhinken (Hartikainen et al., 2022).

Im Feld HCAI ist durch sein Hervorgehen aus der HCI bereits Erfahrungswissen vorhanden, wie Gestaltungsphilosophien erfolgreich für die Anforderungen der Gestaltungspraxis operationalisiert werden können. Besonders der Ansatz der MZG ist hierfür bezeichnend. Jedoch stellen die Mitglieder der HCAI-Community im Kontext des MZKG-Diskurses fest, dass sie zur Realisierung der MZKG-Gestaltungsphilosophie erweiterte Strategien benötigen, wie jene, die bislang in der HCI etabliert sind. Vor diesem Hintergrund kann es sich deshalb für die HCAI-Community lohnen, methodische Vorschläge aus Ansätzen wie CTA und VSD auf ihre Anwendbarkeit zur Realisierung der MZKG zu prüfen. Die in CTA und VSD ausgearbeiteten, übernommenen und modifizierten Methoden sind das Resultat jahrzehntelanger Auseinandersetzungen mit Fragestellungen, die gegenwärtig auch Gegenstand in HCAI-Debatten sind. Die CTA- und VSD-Methoden zeigen darüber hinaus Wege auf, wie erarbeitete theoretische Einsichten methodisch aufbereitet und realisiert werden können. Beispielsweise sind in die Entwicklung der Gestaltungsmethoden von CTA und VSD Kompromisschlüsse eingegangen, um sie in verschiedenen sozialen Kontexten verwenden zu können. Die CTA- und VSD-Methoden unterbreiten Vorschläge, wie ein Balanceakt aus methodischer Pragmatik und Flexibilität erfolgen kann, um in maximal vielen Projektkontexten eine hohe Mensch- und Werteorientierung zu ermöglichen. Diese Anforderungen stellen sowohl die HCAI-Community selbst wie auch Vertreter der Gestaltungspraxis an Methoden der MZKG. Darüber hinaus fordern Gestaltungspraktiker, dass MZKG-Methoden sie dazu befähigen müssen, in lang andauernden Gestaltungsprozessen der Größenordnung von mehreren Jahren anwendbar zu sein, und auch, dass es ihnen mithilfe der Methoden möglich wird, die vielfältigen Anforderungen sehr großer und sehr heterogener Ziel- und Nutzergruppen von KI-Systemen adressieren zu können. Die Methoden müssen darüber hinaus dazu befähigen, die begrenzt verfügbaren Ressourcen effizient und zielführend einzusetzen. Damit fundierte Aussagen getroffen werden können, inwieweit die Methoden von CTA und VSD anschlussfähig zum MZKG-Diskurs in HCAI sind, muss die Praxistauglichkeit der Methoden anhand dieser Kriterien evaluiert werden. Dazu empfiehlt es sich, die Methoden von erfahrenen Gestal-

tungspraktikern aus dem Feld der KI-Gestaltung einschätzen und beurteilen zu lassen, da solche KI-Experten Erfahrung in der Entwicklung und Umsetzung von KI-Anwendungen besitzen und um die besonderen Rahmenbedingungen in KIGP und den Anforderungen darin aktiver Gestalter und Entscheidungsträger wissen. Sie sind aufgrund ihrer Berufserfahrung dazu in der Lage, fachgerechte Aussagen über die Zweckmäßigkeit und den Mehrwert der Gestaltungsmethoden zu treffen, da sie die dafür notwendige Fachkenntnis und Projekterfahrung besitzen.

## 6.2 Vorgehensweise

Für die Untersuchung der Methodik von CTA und VSD führe ich deshalb eine Expertenevaluation durch. Dazu kombiniere ich drei Methoden qualitativer Forschung: (1) die Entwicklung von Storyboards zur inhaltlichen Unterstützung von (2) leitfadengestützten Experteninterviews, sowie (3) die Durchführung einer SWOT-Analyse zur Auswertung der von den Interviewpartnern geäußerten Gedanken und Einschätzungen. Insgesamt befrage ich neun ausgewiesene Fachexperten aus verschiedenen Branchen der praktischen KI-Gestaltung. In den Interviews bitte ich die KI-Experten darum, unter Rückgriff auf ihre langjährige Berufserfahrung in Verantwortungspositionen die Methoden von CTA und VSD speziell im Hinblick auf ihre Eignung für den Kontext von KIGP zu beurteilen. Ziel der explorativen Studie ist es zu ermitteln, inwieweit die Methoden dazu geeignet sind, auch in HCAI zielführende und verlässliche Gestaltungsergebnisse zu fördern. Zur Beantwortung dieser Frage analysiere ich Expertenaussagen mithilfe einer SWOT-Analyse, um die verschiedenen von den KI-Experten zum Ausdruck gebrachten Stärken, Schwächen, Chancen und Risiken der Methoden strukturiert wiedergeben zu können. Die Expertenevaluation der CTA- und VSD-Methoden ergänzt die bislang auf Basis der Rollenkonzepte erarbeiteten Erkenntnisse zur Anschlussfähigkeit von CTA und VSD zu HCAI um den in HCAI sehr bedeutsamen methodischen Aspekt.

### 6.2.1 Vorbereitung

Um die Experteninterviews zielführend durchführen zu können, ist die Durchführung verschiedener vorbereitender Maßnahmen erforderlich. Diese umfassen die Entwicklung von Storyboards, das Gewinnen von Interviewpartnern und das Erstellen eines Interviewleitfadens.



## Entwickeln von Storyboards

Storyboards sind Dokumente die es ermöglichen, durch die Verbindung von Text- und Bildinformationen Konzepte und Ideen kompakt darzustellen und verständlich zu vermitteln (siehe Abb. 45).

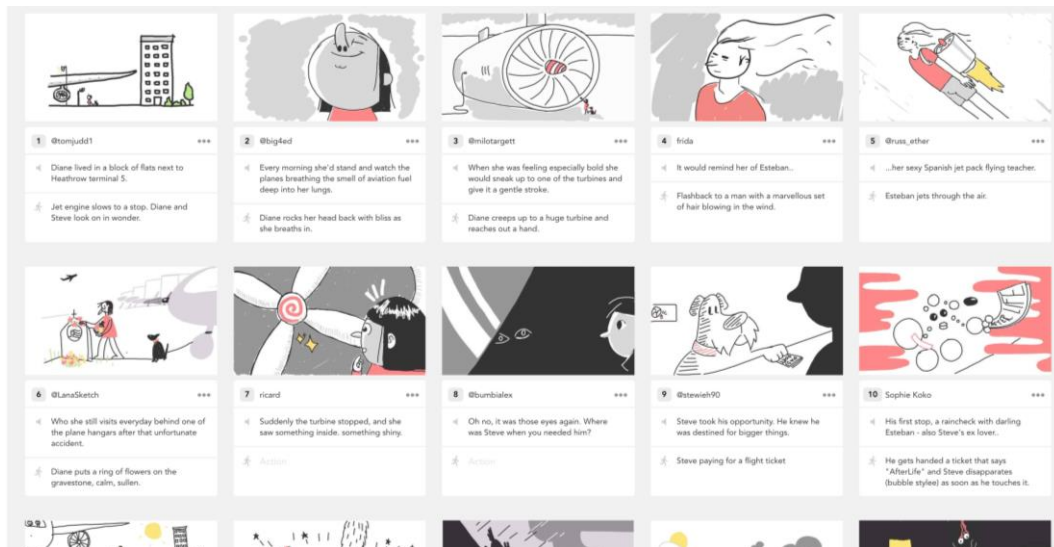


Abbildung 45: Beispiel eines Storyboards (Wiedermann, 2024)

Sie haben ihren Ursprung in der Filmindustrie, wo sie bis heute fester Bestandteil von Drehbüchern sind und zur Produktion von Filmen verwendet werden. Darüber hinaus werden Storyboards auch in weiteren Fachbereichen als Kreativitätstechnik eingesetzt (Storyboarding; Mou et al., 2013; Orr et al., 1994; Truong et al., 2006). So hat sich die Storyboardtechnik beispielsweise in der HCI für die Konzeption und Gestaltung von interaktiven Mensch-Maschine-Schnittstellen als hilfreich erwiesen (Li et al., 2016; Madsen & Aiken, 1993). Für die Durchführung der Expertenevaluation ist die Entwicklung von Storyboards hilfreich, da mir ihre Verwendung ermöglicht, allen befragten KI-Experten einen niederschweligen und verständlichen Zugang zu den CTA- und VSD-Methoden zu gewähren. Sie dienen den KI-Experten als Vorbereitungs- und Referenzdokument, und unterstützen sie bei der Ausführung und Entwicklung ihrer Gedanken. Auch kann ich die Storyboards in den Interviews dazu verwenden, auftretende Verständnisfragen schnell und zufriedenstellend zu beantworten.

Der inhaltliche Aufbau der von mir entwickelten Storyboards folgt einer Struktur, die ich gewählt habe, um den KI-Experten auf einfache Weise Zugang zu relevanten Informationen der CTA- und VSD-Methoden zu ermöglichen. Die finalen Storyboards beinhalten Informationen zu den grundsätzlichen Intentionen von CTA und VSD, detaillierte zusammenhängende Bild- und Texterläuterungen zu den Methoden der beiden Ansätze, sowie zusätzliche Informationen zur Zielsetzung und zum Ablauf der Interviewstudie. Ein Storyboard umfasst insgesamt drei Seiten, bestehend aus einer erläuternden Seite zu Inhalt und Gebrauch des Storyboards, sowie je einer Seite mit erläuternden Informationen zu Metho-

den aus CTA und VSD. Die beiden Seiten zu den Ansätzen beinhalten eine Kurzbeschreibung mit zentralen Informationen zum Ansatz, der die übergeordneten Intentionen des Ansatzes vorstellt und die daraus abgeleiteten Methoden kurz einordnet. Die dafür verwendeten Informationen habe ich der CTA- und VSD-Literatur entnommen. Den Hauptteil einer Ansatzseite eines Storyboards nimmt die konkrete Darstellung und Beschreibung einer Methode in Anspruch, die sich aus Bild- und Textmaterial zusammensetzt. Den größten Teil nehmen dabei die Bildkonstellationen ein, die die Textinformationen zu den Methoden visualisierend verdeutlichen (siehe Abb. 46). Um die Nachvollziehbarkeit der Methoden für die Interviewpartner noch weiter zu erhöhen, habe ich die teilweise komplexen Methodenbeschreibungen aus der Literatur im Sinne einer didaktischen Reduktion auf ihre wesentlichen Elemente reduziert und in mehrere Teilschritte untergliedert.

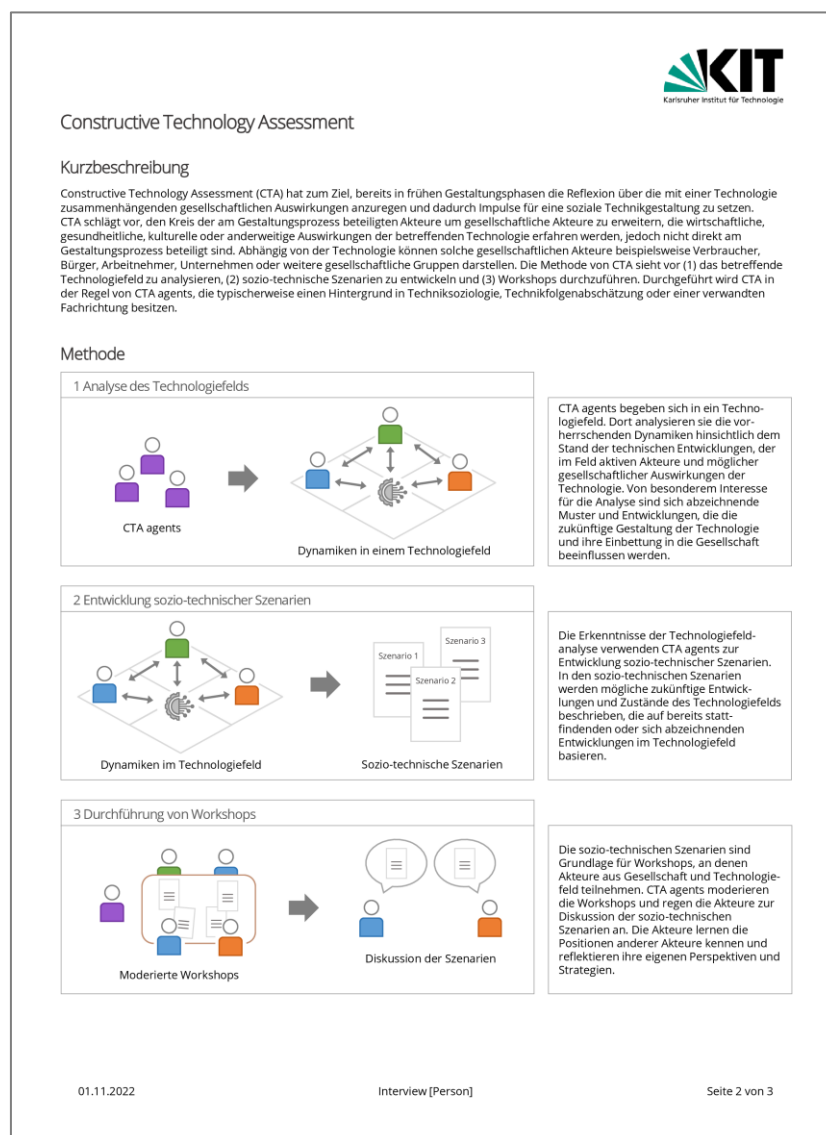


Abbildung 46: Layout einer Ansatzseite in den Storyboards der Interviewstudie

Eine Besonderheit der Entwicklung der Storyboards betrifft die Darstellung der Methoden von VSD. Im Gegensatz zu CTA hält die VSD-Literatur eine deutlich höhere Zahl von Methoden bereit, die in VSD-Projekten angewendet werden können (siehe Kap. 4.3.2). Die zeitliche Begrenzung der Interviews in Kombination mit meinem Bestreben, in jedem Interview die Methoden beider Ansätze beurteilen zu lassen, machte es erforderlich, für die VSD-Seiten der Storyboards eine Auswahl von VSD-Methoden zu treffen, die im Storyboard dargestellt und erläutert werden. Eine Auswahl von VSD-Methoden zu treffen war zudem erforderlich, da eine Beurteilung der tripartite methodology aufgrund ihrer hohen Abstraktionsebene den KI-Experten nur wenig konkrete Aussagen und Beurteilungen ermöglicht hätte. Ich habe deshalb in den Storyboards je eine VSD-Methode dargestellt, die einer der drei Investigations-Typen der tripartite methodology in VSD zugeordnet werden können. Meine Auswahl fiel auf die Methoden (1) der Stakeholderanalyse (repräsentativ für die conceptual investigation), (2) der Value dams and flows (repräsentativ für die empirical investigation), sowie (3) der Value-oriented prototypes (repräsentativ für die technical investigation). Auf diese Weise kann in jedem Interview eine von drei VSD-Methoden durch KI-Experten beurteilt werden. Die Auswahl, in welchem Interview ich welche VSD-Methode zur Beurteilung stelle, erfolgt anhand der Chronologie der Interviewtermine. Da alle VSD-Methoden in der tripartite methodology verankert sind, lassen es die Ergebnisse der Interviewstudie in ihrer Gesamtbetrachtung zu, Rückschlüsse zur Anschlussfähigkeit der tripartite methodology zu HCAI ziehen. Aufgrund der Bedeutung der tripartite methodology in VSD enthalten alle Storyboards entsprechende Informationen zur tripartite methodology, sowie Hinweise, dass in VSD-Projekten (1) idealerweise jede der drei Arten von Untersuchungen durchzuführen sind, und dass (2) für den Zweck des Interviews lediglich eine von mehreren VSD-Methoden als repräsentatives Beispiel dient. Die Storyboards in ihrem vollständigen Umfang sind im Anhang hinterlegt. Sie sind, wie auch der Interviewleitfaden, in deutscher und in englischer Sprache verfasst, da ich potenziellen Studienteilnehmern die Möglichkeit angeboten habe, in beiden Sprachen mit mir zu sprechen.

## Gewinnen von Interviewpartnern

Die Kriterien für eine Teilnahme an der Interviewstudie habe ich bewusst eher weit gefasst, um ein großes Feld potenziell geeigneter Teilnehmer anzusprechen. Die Teilnahmebedingungen der explorativen Interviewstudie lauten deshalb, dass Studienteilnehmer mehrere Jahre Berufserfahrung in der Gestaltung von KI-Lösungen aufweisen müssen, idealerweise in einer Funktion als Gestalter oder als (Projekt-)Manager. Diese Bedingungen habe ich entsprechend kommuniziert, als ich verschiedene Maßnahmen zur Gewinnung von Interviewpartnern ergriff. Einen großen Teil der Studienteilnehmer konnte ich dabei durch die Ansprache meines eigenen beruflichen Netzwerks gewinnen.

Darüber hinaus haben sich einige Interviewpartner im Nachgang eines Fachvortrags zu einer Teilnahme bereit erklärt, den ich im Oktober 2022 auf einer Veranstaltung im Businesskontext gehalten habe. Durch verschiedene Aufrufe in beruflichen sozialen Netzwerken konnte ich zudem Querkontakte herstellen, die in weitere Zusagen zur Teilnahme an der Interviewstudie mündeten.

## Entwickeln eines Interviewleitfadens

Bei der Entwicklung der Fragen des Interviewleitfadens habe ich darauf geachtet, mehrheitlich offene Fragen zu formulieren, um den KI-Experten ein freies Sprechen zu ermöglichen. Die Formulierung der Fragen ist auf eine Weise gewählt, die das Expertenwissen der Befragten aktivieren soll und sie dazu anregen soll, die Eignung der von CTA und VSD vorgeschlagenen Methoden für die Gestaltung von KI-Systemen aus verschiedenen Blickwinkeln einzuschätzen und zu beurteilen. Die KI-Experten sollen verschiedene Anhaltspunkte und Bewertungskriterien anwenden, die für die Fragestellung der Dissertation relevant sind. Im Ergebnis umfasst der Interviewleitfaden eine Einstiegsfrage, in der die Interviewpartner gebeten werden, ihre Expertise und bisherigen Rollen in der KI-Gestaltung kurz zu erläutern. Anschließend sieht der Leitfaden bis zu sechs Fragen vor, die sich auf die in den Storyboards dargestellten Methoden beziehen. Diese erfragen die professionelle Einschätzung der Gesprächspartner, lassen jedoch auch genügend Raum für die Erläuterung von für die KI-Experten relevanten, prägenden oder anderweitig aufschlussreichen Erlebnisse aus ihrem Berufsleben. Am Ende der Interviews wird den KI-Experten durch eine abschließende Frage die Möglichkeit gegeben, weitere aus ihrer Sicht bedeutsame Gedanken zu äußern sowie weitere Erfahrungswerte zu teilen, die im bisherigen Verlauf des Interviews nicht zur Sprache kamen oder die vermeintlich wenig Relevanz für die gestellten Fragen besitzen. Den Abschluss des Leitfadens bildet ein Fragenspeicher, der weitere fünf Fragen umfasst, auf die bei passender Gelegenheit im Verlauf der Interviews zugegriffen werden kann. Insgesamt habe ich darauf geachtet, den Leitfaden möglichst kompakt zu halten, um den teilweise begrenzt zur Verfügung stehenden Zeitkontingenten der Interviewpartner Rechnung tragen zu können. Der Interviewleitfaden ist im Anhang hinterlegt.

### 6.2.2 Durchführung

In der Interviewstudie befrage ich neun KI-Experten, die alle über mehrere Jahre Berufserfahrung in der KI-Gestaltung verfügen. Das Teilnehmerfeld (siehe Tab. 5) besteht aus zwei Frauen und sieben Männern, die in ihren Berufen sowohl für die technische Gestaltung von KI-Systemen verantwortlich sind, wie auch für das Treffen von Managemententscheidungen. Durchgeführt habe ich die Interviews im Zeitraum von November 2022 bis Januar 2023. Zu diesem Zeitpunkt sind die befragten KI-Experten in verschiedenen Unternehmen mit Sitz in Deutschland tätig. Ihre konkreten Positionsbezeichnungen in diesen Unternehmen variieren stark, so gehören zum Teilnehmerfeld beispielsweise Senior Research Scientists, CEOs oder auch Product Owner für Conversational AI.

Interview Nr.	Positionsbezeichnung	Branche
1	Senior Research Scientist, Automatic Speech Recognition	Softwareentwicklung
2	Co-Founder & Managing Director	Industrielle Fertigung
3	Senior Teamleader Artificial Intelligence	Robotik
4	Fachbereichsleitung Corporate Technology	IT-Sicherheit
5	Head of AI & Data Science, Public Sector & Defense	IT & Unternehmensberatung
6	Co-Founder & Managing Director	Softwareentwicklung
7	Product Owner Conversational AI	Softwareentwicklung
8	Senior Manager AI Business Consulting	Lebensmitteleinzelhandel
9	CEO & Data Scientist	IT & Unternehmensberatung

Tabelle 5: Teilnehmerfeld der Interviewstudie

Sieben Experten entscheiden sich für die Onlinedurchführung des Interviews, ein Interview wird in den Räumen des Instituts für Technikfolgenabschätzung und Systemanalyse Karlsruhe durchgeführt und ein Interview findet in Räumen des Unternehmens statt, für den der befragte KI-Experte tätig ist. Die Interviews sind immer 1:1-Gespräche und dauern zwischen 30 Minuten und 60 Minuten. Ein Interview wird auf Englisch geführt, acht auf Deutsch. Zu Auswertungszwecken fertige ich von allen Interviews Tonaufzeichnungen an.

Der Ablauf aller Interviews orientiert sich an der Reihenfolge der Fragen im Interviewleitfaden, stellenweise variiert jedoch die Reihenfolge der Fragestellungen zu den CTA- und VSD-Methoden, abhängig von den konkreten Äußerungen und Gedankengängen, die die Experten verbalisieren. In den Interviews achte ich darauf, den Experten das Fortführen ihrer Gedankengänge möglichst einfach zu machen. Bereits vor den Interviewterminen stelle ich allen Studienteilnehmern ein Storyboard zur Verfügung, verbunden mit der Bitte, sich vor dem Interview mit den Inhalten des Storyboards auseinanderzusetzen. Zu Beginn der Interviews bitte ich die Experten, ihre Expertise im Bereich der KI-Gestaltung darzulegen (Einstiegsfrage). Anschließend erfolgt durch mich eine kurze Erläuterung der Storyboardinhalte, um Missverständnissen und Fehlinterpretationen vorzubeugen und etwaige Rückfragen zu beantworten. Der Hauptteil der Interviews ist den inhaltlichen Fragen zu den CTA- und VSD-Methoden gewidmet. Ich bitte die Experten, die für sie hervortretenden Auffälligkeiten und Merkmale der Methoden zu benennen, und zu erläutern, weshalb diese Merkmale für sie von Bedeutung sind. Die Experten sollen anschließend möglichst konkrete Argumente anführen, weshalb sie die dargestellten Methoden für geeignet oder ungeeignet befinden, um sie in der Gestaltung von KI-Systemen einzusetzen. Dazu fordere ich die Experten auch explizit auf, ihrer professionellen Meinung Ausdruck zu verleihen

(Was begrüßen Sie? Wo sehen Sie Lücken?) und entsprechend ihr Erfahrungswissen zu aktivieren. Um einen Perspektivwechsel zu vollziehen, bitte ich die Experten gegen Ende des Interviews darum, sich vorzustellen, sie würden in einem gegenwärtigen oder zukünftigen Projekt aktiv mit den dargestellten CTA- und VSD-Methoden arbeiten. Ausgehend von dieser Vorstellung sollen die Experten ihre bisher angebrachten Gedanken und Argumente überdenken und erhalten dadurch auch die Möglichkeit, Änderungen oder Ergänzungen an ihren bisherigen Darlegungen vorzunehmen.

Da ich die Storyboards als wesentliches Referenzdokument für die Experten konzipiert habe, habe ich sie den Experten auch während der gesamten Dauer der Interviews zur Verfügung gestellt. In den Onlineinterviews wird dies durch die Screensharingfunktion des verwendeten Conferencingtools ermöglicht, in den Offlineinterviews habe ich den Experten die Storyboards zusätzlich in Printform zur Verfügung gestellt. Im Verlauf der Interviews habe ich ferner darauf geachtet, dass den Methoden beider Ansätze das gleiche Maß an Aufmerksamkeit gewidmet wird. Dies macht es erforderlich, die Experten stellenweise auf die CTA- oder VSD-Methode zurückzuführen und konkrete Gedanken zu ihnen einzufordern. Zusätzlich habe ich allen Studienteilnehmern Anonymität zugesichert.

### 6.2.3 Auswertung

Die KI-Experten benennen an den CTA- und VSD-Methoden verschiedene begrüßenswerte Aspekte, äußern jedoch auch begründete Kritik. Darüber hinaus identifizieren die Experten verschiedene positive wie auch negative Potenziale, die die Methoden für den Einsatz zur Gestaltung von KI-Systemen bereithalten. Um die genannten Merkmale und Argumente systematisch erfassen und analysieren zu können, führe ich für die Auswertung der Experteninterviews eine SWOT-Analyse (z.B. Ghazinoory et al., 2011; Gurl, 2017; Helms & Nixon, 2010) durch (siehe Tab. 6).

	Strenghts (Stärken)	Weaknesses (Schwächen)
Opportunities (Chancen)	S-O-Analyse	W-O-Analyse
Threats (Risiken)	S-T-Analyse	W-T-Analyse

Tabelle 6: Eine typische SWOT-Matrix (Skinner et al., 2012, S. 163)

Die SWOT-Analyse ermöglicht es mir, die verschiedenen von den Experten identifizierten Stärken, Schwächen, Chancen und Risiken der Methoden auf strukturierte Weise zu erfassen und wiederzugeben. Sie eröffnet fundierte und differenzierte Perspektiven auf die Methoden der beiden Ansätze und ermöglicht eine realistische Einschätzung der Beitragfähigkeit zu Methoden-Fragestellung zur MZKG in HCAI. Datengrundlage der SWOT-Analyse stellen die Audioaufzeichnungen der Interviews dar, die während den Interviews angefertigten Notizen, sowie die Gedächtnisprotokolle, die ich im Anschluss an die Experteninterviews erstellt habe. Die für die Fragestellung der Dissertation besonders wichtigen Aussagen der KI-Experten transkribiere ich im Rahmen der SWOT-Analyse. Trotz der Einordnung der Aussagen in Stärken, Schwächen, Chancen und Risiken sind die Aussagen immer in ihrer zusammenhängenden Bedeutung zu interpretieren.

## 6.3 Stärken

Die Stärken beschreiben die Elemente der Methoden von CTA und VSD, die die KI-Experten in Bezug auf die menschenzentrierte Gestaltung von KI-Systemen begrüßen, und über deren wertvollen und zielführenden Beitrag zu Gestaltungsprozessen sie Gewissheit zum Ausdruck bringen. Grundsätzlich sind die KI-Experten in der Lage, die übergeordneten Ziele der beiden Ansätze nachzuvollziehen und erkennen den Beitrag, den die vorgeschlagenen Methoden zum Erreichen dieser Ziele zu leisten anstreben. Die verschiedenen von den Experten angeführten Stärken werde ich nachfolgend, sowie für die nachfolgenden Kapitel, thematisch gruppiert wiedergeben. Stellenweise beziehen sich die Experten dabei auf konkrete methodische Aspekte, stellenweise jedoch auch auf die übergeordneten Ideen und Vorschläge von CTA und VSD.

### 6.3.1 Prozess- und Ergebnisorientierung

In den Interviews kommt verschiedentlich zum Ausdruck, dass die in CTA und VSD vorgeschlagenen Methoden strukturierte und nachvollziehbare Gestaltungsmaßnahmen vorsehen, weshalb sie sich grundsätzlich für die Integration bzw. Anwendung in KIGP eignen würden. Konkret befindet ein Experte die Methoden für tauglich, um die bei der Gestaltung von KI-Systemen auftretende Probleme zu analysieren und zu lösen, insbesondere, wenn in Unternehmen nur wenig methodisches Knowhow vorhanden ist:

*„Es methodisch anzugehen [...] wenn man wirklich wenig Knowhow im Unternehmen hat, dann jemanden einzustellen der [...] schaut wo es Probleme gibt, die mit aktuellen Entwicklungen aus der Forschung gelöst werden können, ist mit Sicherheit ein richtiger Ansatz“ (Interview 2, Min. 11:48).*

Dies ist nach Angaben des Experten insbesondere in Hinblick auf die systematische Auseinandersetzung mit potenziellen Chancen und Risiken von KI-Systemen wertvoll und notwendig, da nach seiner Erfahrung Fragen zur gesellschaftlichen Akzeptanz eines KI-Systems einen KIGP durchgehend begleiten. Diesbezüglich berichtet der Experte von immer wieder auftretenden Problemen, da diesem Aspekt in Gestaltungsprozessen zu

wenig Aufmerksamkeit gewidmet wird. Die Anwendung der CTA- und VSD-Methoden helfe Gestaltern und Entscheidungsträgern nicht nur, frühzeitig mögliche Akzeptanzprobleme zu identifizieren und adressieren zu können, sondern auch die Erwartungshaltungen und Annahmen von Menschen zu den Funktionen eines KI-Systems von häufig überzeichneten Vorstellungen auf realistisch mögliche Funktionen zurückzuführen. Einen weiteren Mehrwert bescheiden die Mehrzahl der Experten den Methoden dahingehend, dass sie eine klare Zuweisung von Verantwortlichkeiten ermöglichen. Dieser Aspekt besäße für KIGP in der Praxis eine besondere Bedeutung, da erst durch transparent kommunizierbare Handlungsabläufe klare Verantwortungsbereiche und konkrete Tätigkeiten im Gestaltungsgeschehen abgestimmt werden können. Dies trage sehr stark dazu bei, eine realistische Chance auf Gestaltungserfolg zu kreieren und zu wahren. Die CTA- und VSD-Methoden seien insofern nachvollziehbar, da sie strukturierte Handlungsabläufe beschreiben, die sowohl inhaltlich verständlich wie auch organisatorisch realisierbar seien. Unter diesem Gesichtspunkt kommt in den Interviews jedoch auch zum Ausdruck, dass die Methoden von VSD tendenziell als geeigneter eingeschätzt werden, Verantwortungsbereiche zu definieren und konkrete Handlungsanweisungen abzuleiten, wie die Methoden von CTA:

*„Ich finde das [VSD] pragmatischer, da kann ich besser Handlungsanweisungen ableiten. Das kann ich tatsächlich, wie man so schön sagt, besser operationalisieren in der Praxis“ (Interview 5, Min 22:50).*

Dies bezieht sich auch auf den Gesichtspunkt, dass die CTA- und VSD-Methoden dazu befähigen würden, Zwischenziele zu definieren. Anhand von Zwischenzielen, so die Äußerungen des KI-Experten aus Interview 3, können die Teilfortschritte sowie der Gesamtfortschritt eines Gestaltungsprojekts nachverfolgt werden, was in der Gestaltungspraxis eine zentrale Notwendigkeit darstelle. Der Experte führt aus, dass seine Position als Teamleiter es oftmals erfordere, den tagesaktuellen Fortschritt eines Gestaltungsprojekts in Erfahrung zu bringen, um diesen beispielsweise mit anderen Abteilungen oder Geschäftspartnern diskutieren zu können, oder um notwendige Änderungen an gesetzten Zwischenzielen vorzunehmen. Das Formulieren von Zwischenzielen helfe dabei, für alle Beteiligten die Komplexität des Gestaltungsvorhabens von Anfang an zu reduzieren, da es durch eine Aufteilung in kleinere Zwischenziele verständlicher und realistischer wird. Maßgeblichen Beitrag der CTA- und VSD-Methoden zur Definition und Überprüfung von Zwischenzielen sieht der Experte in den Bestrebungen der Methoden, Deliverables im Sinne von Endprodukten und Endergebnissen zu generieren:

*„Die Strukturiertheit [finde ich gut] und dass man Deliverables hat. Also da kommt etwas rum wie z.B. die Szenarien, die sind ja da, die kann man nachlesen“ (Interview 3, Min 22:58).*



Die leichte Favorisierung der VSD-Methoden kommt auch in Interview 8 zur Sprache. In diesem Interview begründet der KI-Experte seine leicht favorisierende VSD-Tendenz mit dem Argument, dass die Methode ihn in der Wahrnehmung seiner Aufgabe als Projektmanager unterstütze und sich gut in die Durchführung eines Gestaltungsprojekts integrieren lasse:

*„Für mich fühlt sich [VSD] praktischer an, weil es gut in so ein Projektmanagementkontext reinpasst. Wenn man gut leitet, macht man auf jeden Fall auch eine Stakeholderanalyse“* (Interview 8, Min 16:20).

*„Ich glaube das kann man sehr gut in einen Projektrahmen integrieren“* (Interview 8, Min 21:00).

Er bezieht sich in seiner Argumentation auf seine Erfahrung als Führungskraft, aus der er den Wert einer möglichst intensiven Auseinandersetzung mit identifizierten und potenziellen Stakeholdern kennt. Auch deutet er an, dass ihm die Anwendung von Methoden bereits bekannt ist, die ähnliche Ziele verfolgen wie die in VSD vorgeschlagene Stakeholderanalyse. Tatsächlich ist dies ein Aspekt, der sich auch in weiteren Interviews positiv auf die Beurteilungen der Methoden auswirkt. Die Bekanntheitsgrade und Erfahrungswerte, die einige der KI-Experten bereits in der Vergangenheit mit der Anwendung ähnlicher Methoden in der Vergangenheit erzielt haben, wirken sich befürwortend für die Anwendung der CTA- und VSD-Methoden aus. So setzte sich ein Experte sehr fürsprechend für die Verwendung der Fragebögen aus, die die Methode der Value dams and flows vorschlägt. Laut Aussage des Experten wird in dem Unternehmen, für das er zum Zeitpunkt des Interviews tätig ist, bereits mit ähnlichen Fragebögen gearbeitet, um die Risiken von KI-Lösungen zu identifizieren und sie zu beurteilen:

*„Solche Fragebögen zu Risiken die existieren bereits, bzw. wir haben auch tatsächlich sowas zu bestimmten Aspekten von KI [...]. Deshalb würde ich [VSD] als pragmatischer empfinden“* (Interview 5, Min 25:35).

Die Anwendung der (gedanklichen) Metapher von Dämmen und Flüssen im Methodenkontext ist neu für den Experten, sie erschließt sich ihm im Zusammenhang mit Wertefragen jedoch schnell. Er begrüßt dabei, dass „Wertedämme“ als zentrale Risikokriterien in KI-Gestaltungsprozessen berücksichtigt werden können, auch wenn ihre Identifikation erst zu fortgeschrittenen Zeitpunkten erfolgt. Auf ähnliche Weise wird in einem anderen Interview die in VSD vorgeschlagene Methode der Value-oriented prototypes für zielführend befunden. Der in Interview 3 befragte KI-Experte hat in der Vergangenheit mit Methoden gearbeitet, die ähnliche Vorgehensweisen vorsehen, um angedachte technische Funktionen niederschwellig und partizipativ zu realisieren, um von Stakeholdern belastbares Feedback über die angedachten Funktionen einzuholen. Mit Verweis auf diese Erfahrungen bringt er im Interview zum Ausdruck, dass sich die Value-oriented prototypes besonders gut dazu eignen würden, in Gestaltungsprozessen, in denen bereits konkrete Einsatzszenarien von KI-Systemen identifiziert wurden, die KI-Systeme oder einzelne ihrer Funktionen unter ethischen Gesichtspunkten zu analysieren und von Stakeholdern

bewerten zu lassen. Für ihn besteht der besondere Mehrwert der Methode darin, dass sie als Ergebnis die Gestaltung konkreter, materieller Artefakte vorsieht:

*„Klar, das macht total Sinn, dann hat man etwas Handfestes [...]. Für konkrete Produkte halte ich das auf jeden Fall sinnvoll“* (Interview 3, Min 24:10).

Als eine weitere Stärke der Methoden identifiziert der KI-Experte in Interview 7, dass sie dabei helfen, den Rahmen eines Gestaltungsprojekts zu definieren:

*„Ich glaube da kann man je nach Projekt entscheiden, wie tief man dann da reingehen möchte oder wie weit man das Ganze fassen möchte. Bei den indirekten Stakeholdern kann ich mir z.B. gut vorstellen, auf wie viele Level man das bringen [kann] oder auf wie viele Dimensionen von Stakeholdern man das bringen [kann]. [...] Ich muss sagen ich finde [KI] eigentlich einen perfekten Use Case. Von den Ansätzen und den Ideen [...] finde ich passt es sehr sehr gut in dem Rahmen“* (Interview 7, Min 39:12).

Er bringt zur Sprache, dass sowohl die CTA- wie auch die VSD-Methode es ihren Anwendern erlauben würden, zu steuern, wie viel Zeit und Energie sie sowohl ihrer Durchführung wie auch der Auseinandersetzung der mit ihrer Hilfe hervorgebrachten Erkenntnisse widmen. Übertragen auf die konkreten Methoden spricht er damit die Entscheidungsmöglichkeit an, beispielsweise bei der Durchführung der CTA-Methoden zu steuern, wie ausführlich ein Technologiefeld in CTA analysiert wird (z.B. durch die Definition der Anzahl von Organisationen, die besucht werden), oder wie viele soziotechnische Szenarien anschließend entwickelt werden. Ihm gefällt die Möglichkeit, die konkreten Rahmenbedingungen und den „Aufwandsumfang“ der Methoden an die Notwendigkeiten und Möglichkeiten eines konkreten Gestaltungsprojekts anzupassen. Er befindet die organisatorische und inhaltliche Flexibilität der Methoden insbesondere vor dem Hintergrund für hilfreich, da es in KI-Projekten häufig sehr wichtig sei, schnell auf aktuelle Entwicklungen reagieren zu können und sich somit eine gewisse Zeit- und Handlungsflexibilität zu bewahren. Umgekehrt könne, wenn die Methoden zur Identifikation zentraler Gestaltungsfaktoren beitragen, weitere Ressourcen aufgewendet werden um sich beispielsweise ihrer vertieften Auseinandersetzung zu widmen, oder um Maßnahmen zu entwickeln, die ihre Berücksichtigung im Gestaltungsgeschehen ermöglichen.

### 6.3.2 Befähigung zu Reflexion und Innovation

Eine weitere Stärke schreiben die KI-Experten den beiden Ansätzen dahingehend zu, dass sie sämtliche am Gestaltungsprozess beteiligten Personen zu einer weitreichenden Reflexion über das Gestaltungsgeschehen und die Gestaltungsinhalte befähigen. Alle Experten wissen aus Erfahrung, dass KIGP neben vielerlei technischen und organisatorischen Aspekten stets auch Fragestellungen zu sozialen und ethischen Inhalten berühren, die jedoch in der Gestaltungspraxis aufgrund von Ressourcenknappheit und dem Fokus auf Wirtschaftlichkeit häufig nur in geringem Umfang adressiert werden können. Deshalb begrüßen viele der KI-Experten die durch die Methoden operationalisierte Intention von CTA und

VSD, insbesondere Gestaltern und Entscheidungsträgern Möglichkeiten zur Auseinandersetzung mit sozialen und ethischen Fragestellungen an die Hand zu geben. Die Wertschätzung für derlei Bestrebungen in den beiden Ansätzen und ihren Methoden kommt dabei auf verschiedene Weise zum Ausdruck:

*„Ich finde es sehr positiv, dass man versucht so ein Thema ganzheitlich anzugehen und sich nicht nur auf die technische Umsetzung fokussiert, auf die harten technischen Anforderungen, [...] sondern [sich] auch [fragt]: Was erreiche ich eigentlich mit dieser Lösung?“ (Interview 7, Min 25:08).*

*„Wenn ich einen [Value] dam identifiziere, dann kann ich das als Bedingung in der KI-Entwicklung berücksichtigen. Oder den Optimierungsprozess ganz anders aufsetzen, dass ich auf ganz andere Aspekte in der Input-Domäne abstelle. Es geht auch in die umgekehrte Richtung, wenn eine Folge als besonders wünschenswert identifiziert ist, dann kann man die Entwicklung dahingehend steuern, dahingehend lenken. Von daher finde ich [VSD] viel viel nützlicher“ (Interview 5, Min 26:40).*

*„Man hat dann so ein breites Gebiet an Stakeholdern, dass da so ein Workshop, wo die alle sich mal an einen Tisch setzen und ihre Seite darstellen und sich Gedanken machen, genau richtig [ist]“ (Interview 4, Min 22:24).*

Die KI-Experten machen deutlich, dass sie die vorgeschlagenen Methoden für adäquat gewählte bzw. geschaffene Werkzeuge halten, um Gestalter und Entscheidungsträger mit ihren Verpflichtungen in sozialen und ethischen Tätigkeitsbereichen und Auswirkungsfeldern zu konfrontieren und sie darüber hinaus zur Lösungs- und Kompromissfindung entsprechender Herausforderungen befähigen. Während manche der KI-Experten sich in ihren Ausführungen auf mehr übergeordnete Aussagen zur Reflexionsbefähigung beschränken, benennen andere KI-Experten konkrete Eigenschaften und Vorkehrungen der Methoden. So begrüßt der KI-Experte in Interview 4 explizit, dass die Methoden zur Identifikation nicht-intendierter Effekte befähigen:

*„Es ist gut das hier (deutet auf Schritt 2 im Storyboard, die Identifikation möglicher nicht-intendierter Nutzungskontexte und Folgen) explizit dastehen zu haben, da bin ich einverstanden. [...] Aber ich glaube die andere Methode, dass man sich die möglichen Zukunftsszenarien überlegt, bringt einen schon in diese Denkweise, out of the box zu denken und sich zu überlegen, was kann denn passieren, womit ich jetzt nicht gerechnet habe“ (Interview 4, Min 28:50).*

Wie aus dieser Aussage deutlich wird, befindet der KI-Experte in Interview 4 die CTA-Methoden im Vergleich mit den VSD-Methoden für besser geeignet, um die an KIGP beteiligten Gruppen und Individuen zur Reflexion zu befähigen. Dies begründet der Experte anschließend mit Verweis auf die vorgeschlagene Vorgehensweise der CTA-Workshops, die die Diskussion der entwickelten soziotechnischen Szenarien im Verbund verschiedener Akteure des Technologiefelds vorsieht:

*„Es ist halt die Frage, wenn sich eine Person im Kämmerle ihr Zukunftsszenario ausdenkt – da kommen wahrscheinlich verschiedene Szenarien heraus, aber [...] im Austausch ist es dann besser, weil man nochmal Input kriegt und nochmal auf Sachen kommt. Deshalb macht das eigentlich nur im Austausch Sinn“ (Interview 4, Min 29:20).*

Die von diesem KI-Experten angesprochene Funktion der Methoden, in Gestaltungsprozessen involvierte Personen zu einem out of the Box-Denken zu bewegen, wird auch von anderen KI-Experten begrüßt. Sie begrüßen die Eigenschaft der Methoden, die Kreativität aller Beteiligten zu fördern, indem sie die Beteiligten dazu auffordern, sich mit nicht-alltäglichen, aber dennoch relevanten Fragestellungen und Szenarien auseinanderzusetzen. Das kreative Denken in KIGP möglichst kontinuierlich aufrecht zu erhalten und immer wieder neu zu anzustoßen sei von großer Bedeutung, da auf diese Weise das angestrebte Schaffen von technischen, interaktiven und auch sozialen Innovationen unterstützt werden könne. Die CTA- und VSD-Methoden seien in der Lage, diese Bestrebungen zu unterstützen und darüber hinaus dazu beizutragen, dass der Gestaltungsprozess den Beteiligten auch Vergnügen bereite, wie es der KI-Experte in Interview 6 formuliert:

*„Das mit dem Prototyping finde ich persönlich viel amüsanter [...]. Warum? Ich möchte wissen, welche Begrifflichkeiten, wie ihre Vorstellung aussieht [...]. Plötzlich merkt man, dass die Leute dann sagen „ich stelle mir vor, also wie wäre es, wenn das, das oder das passiert“. Das finde ich super. [...] Dann kriegt man Ideen [...] wie Menschen Probleme einfach lösen können [...] oder auch selber inspiriert werden. Diese Inspiration ist eine tolle Quelle erstens um den Kunden, oder die Gesellschaft auch an dieser Stelle zu verstehen und zweitens, dass man sich näherkommt und in einer Interaktion über dieses Problem diskutiert“ (Interview 6, Minute 18:35).*

Dieser Experte identifiziert demnach eine Stärke der Value-oriented Prototyping-Methode aus VSD, fachlich wertvolle Erkenntnisse auf interessante und kognitiv wenig anspruchsvolle Weise zu erarbeiten, und gleichzeitig mehr über die Denkweisen und Problemlösestrategien von Stakeholdern zu erfahren. Der Experte schätzt sehr, dass die Methode vorsieht, sämtlichen Beteiligten eine gemeinsame und interaktive Reflexion sowie einen kollaborativen Wissenserwerb zu ermöglichen, um eine menschenzentrierte Gestaltung zu fördern. Mit der Befähigung zur Reflexion verbindet ein KI-Experte eine weitere Stärke der Methoden. Er spricht in Interview 1 davon, dass die Methoden ihre Anwender durch die Antizipation möglicher Folgen dabei unterstütze, Ressourcen ziel führend einzusetzen und Fehlinvestitionen vorzubeugen:

*„They may serve for saving money for companies in the sense that if they anticipate what are possible issues that may occur later on, then we will be working from the beginning with these features. [...] If you identify some requirements for the system early stage, it is much easier to evolve the system from there than from already having a running program“ (Interview 1, Min 20:47).*

Der Experte sieht eine Stärke der Methoden darin, insbesondere in frühen Phasen von Gestaltungsphasen relevante Gestaltungsfaktoren zu identifizieren, diese im weiteren Verlauf zu berücksichtigen und dadurch Mehrkosten zu vermeiden, die durch tiefgreifende Anpassungsmaßnahmen entstehen können. Er bezieht sich im Verlauf des Interviews auf seine Erfahrungen aus Projekten der Vergangenheit, wo Features von KI-Systemen rückwirkend versucht wurden, zu korrigieren und anzupassen. Im Verlauf solcher Korrekturvorgänge werde häufig deutlich, dass Korrekturmaßnahmen durch das Erfassen weniger zusätzlicher Informationen zu Beginn eines Projekts hätten frühzeitiger erkannt werden können. Der KI-Experte begrüßt deshalb den Vorschlag der CTA- und VSD-Methoden, im Rahmen von beispielsweise der Anforderungserhebung eines KI-Systems stärker auf soziale und ethische Aspekte zu achten und dafür eine geringfügige Anstrengungserweiterung zu Beginn eines Gestaltungsprozesses in Kauf zu nehmen.

Für den KI-Experten in Interview 8 steht auch fest, dass die Befähigung zur Reflexion eine Stärke der CTA- und VSD-Methoden darstellt. Beginnend auf einer etwas höheren Argumentationsebene erläutert der KI-Experte, dass die Methoden Gestalter und Entscheidungsträger zu einer Auseinandersetzung mit mehr grundsätzlichen Fragestellungen befähigen, die der Experte selbst im Zusammenhang mit der Gestaltung von KI-Systemen als besonders erforderlich erachtet. Er orientiert sich in seiner Begründung besonders an der CTA-Methode:

*„Ich glaube da ist eine solche Methode ganz gut geeignet, dass man sich ein wenig löst von dem konkreten Projekt, vom konkreten kleinen Geschäftsprozess. Da die Feinheiten auszuarbeiten und sich mal grundsätzlich zu überlegen, was kann denn passieren [...] wenn so eine KI [z.B.] in der Preisgestaltung eingesetzt wird, was kann gut laufen, was kann weniger gut laufen“ (Interview 8, Min 24:20).*

Er begrüßt, dass die CTA-Methode erlaubt, im Projektteam sowie im Austausch mit Stakeholdern gedanklich höhere Abstraktionsebenen einzunehmen, und gemeinsam zu überlegen, welche gesellschaftlichen Auswirkungen der Einsatz von KI-Systemen in bestimmten Lebensbereichen haben kann. Er führt anhand des von ihm gewählten Beispiels der KI-gesteuerten Preisgestaltung aus, dass dies verschiedene Auswirkungen für die Unternehmens- und Kundenseite hätte, und dass in Branchen wie dem Verkauf von Lebensmitteln eine direkte und hohe gesellschaftliche Relevanz gegeben sei. Deshalb sei es ratsam, sich zuvor intensiv mit möglichen Auswirkungen auseinanderzusetzen. Aussagen des Experten aus Interview 7 unterstützen seine Einschätzung, dass die Methoden von CTA und VSD für die Gestaltung von KI-Systemen unter dem Reflexionsaspekt sehr gut geeignet seien:

*„Ich finde, dass die Ansätze hier absolut zu KI-Projekten passen, sogar extrem gut, weil ich hier mit diesem Aspekt [arbeite] – Was bewirke ich eigentlich mit dieser Lösung die ich da technisch versuche umzusetzen? Was bedeutet es, wenn man es jetzt einmal bezieht auf meine Zielgruppe [...], was passiert dann, wenn solch eine Technologie eingesetzt wird?“ (Interview 7, Min 34:46).*

Auch dieser KI-Experte befindet die Methoden für dazu geeignet, sich mit den mehr übergreifenden ethischen und gesellschaftlichen Fragestellungen zu befassen, die im Zusammenhang der Gestaltung von KI-Systemen in den Vordergrund treten. Eine intensivere Auseinandersetzung mit möglichen gesellschaftlichen Folgewirkungen von KI-Systemen begrüßt er, und führt für seine Position das Beispiel einer durch KI geförderten Produktionssteigerung an, die er in wenigen Sätzen gedanklich entwickelt. Wenn Menschen mit Hilfe von KI künftig produktiver würden, könne das beispielsweise zur Verschiebung von Aufgabenbereichen innerhalb bestimmter Organisationseinheiten führen. Darin sieht der Experte wiederum ein Potenzial für Unzufriedenheit, da eine mögliche Gefahr darin bestünde, Expertise und Kompetenzen zu schnell zu verlagern. Für die Auseinandersetzung mit solchen aus seiner Sicht realistischen zukünftigen Entwicklungen befindet er die Methoden für hilfreich und zielführend:

*„Ich finde, dass beide [Ansätze] wichtig sind und auch extrem interessant um die Auswirkungen abschätzen zu können“ (Interview 7, Min 35:40).*

### 6.3.3 Partizipation von Stakeholdern

Die KI-Experten schätzen an den Methoden beider Ansätze sehr die verschiedenen Eigenschaften und Vorgehensweisen, die es ermöglichen, verschiedene Stakeholder aktiv in den Gestaltungsprozess zu integrieren und ihnen Möglichkeiten anzubieten, zum Gestaltungsgeschehen beizutragen. Sämtliche der befragten Experten haben Erfahrungen in der Anwendung partizipativer Methoden, entweder aus technischen Kontexten oder aber aus Projekten, die sich Aspekten wie beispielsweise Kundenzufriedenheit widmen. Aus dieser Erfahrung schöpfen sie, wenn sie die partizipativen Elemente der CTA- und VSD-Methoden beurteilen und begrüßen insbesondere die Schwerpunktsetzung zur Ermittlung und Beteiligung potenziell negativ betroffener Gruppen und Individuen, wie beispielsweise in Interview 3 deutlich wird:

*„Was sehr gut ist, ist, dass die nicht-technischen Stakeholder auch mit dabei sind. Da [kenne ich] einen tollen Spruch [aus dem Kontext der Arbeit] mit Behinderten: „Nicht über uns ohne uns.“ [...] Und das macht ja auch Sinn“ (Interview 3, Min 23:16).*

Solche Gruppen stärker an Prozessen der KI-Gestaltung zu beteiligen halten sie für sinnvoll und zielführend, da mit KI-Projekten häufig eine höhere Zahl an unbekannten Faktoren und Auswirkungen verbunden sei. Ein sehr zentrales Argument benennen sie ferner mit der Unterstützung von Gestaltern und Entscheidungsträgern, durch die Methoden in direkten Kontakt mit Stakeholdern zu treten, und durch Interaktion mit Stakeholdern die

Denkweisen und Begrifflichkeiten von Stakeholdern zu ergründen und nachzuvollziehen. Dies hat insbesondere für den KI-Experten in Interview 6 eine hohe Relevanz, da in der von ihm vertretenen Unternehmens- und Gestaltungsphilosophie die starke Orientierung an konkreten Kundenbedürfnissen eine hohe Priorität erfährt. Durch seine Erfolge in der Umsetzung dieser Philosophie in der Vergangenheit sieht er sich bestätigt, weshalb für ihm der Mehrwert partizipativer Elemente in Gestaltungsprozessen von KI-Systemen recht naheliegt:

*„Ich finde [...] wie das durchgeführt wird durchaus relevant [...]. Vor allem die [VSD-]Methode ist, weil man sich das plastisch vorstellen kann [...] interaktiver. Interaktion [mit Stakeholdern] finde ich persönlich immer – das ist eine Präferenzsache – schöner (Interview 6, Min 25:48).*

Daran anknüpfend hebt der Experte in Interview 8 mit dem Aufbau von Empathie eine weitere, für ihn besondere Qualität der Methoden von CTA und VSD hervor. Er führt aus, dass mit der Beteiligung „wirklich Betroffener“ (Interview 8, Min 30:06) ein besonderer „Impact“ verbunden sei, da sich Gestalter und Entscheidungsträger mit sehr konkreten Folgen in sehr konkreten Menschenleben auseinandersetzen können (oder müssen). Dadurch können Gestalter und Entscheidungsträger einen stärkeren Realitätsbezug herstellen, der durch den Austausch mit ihren Mitmenschen intensiviert wird und aus dem sie Rückschlüsse für ihre Gestaltungshandlungen sowie auch für gesamte Prozessverläufe ziehen können:

*„Damit [hat man] so eine Art Empathieansatz. Dann ist vielleicht der Produktverantwortliche konfrontiert mit jemandem, der negativ betroffen ist von dieser Entwicklung. [...] Dann hat man den direkten Kontakt, dann entsteht da empathischer Austausch, und vielleicht auch eine Idee, wie man in der Verbindung mit dem Mensch nochmal eine bessere Lösung findet (Interview 8, Min 31:26).*

Viele der KI-Experten bringen zwischen den Zeilen zum Ausdruck, dass eine menschenzentrierte KI-Gestaltung für sie viele Partizipationsbemühungen bedeutet und voraussetzt, dass sich Gestalter und Entscheidungsträger in möglichst regelmäßigen Abständen mit konkreten Bedürfnissen und Herausforderungen von direkt und indirekt betroffenen Stakeholdern auseinandersetzen. Die Wahl beider Ansätze, für die Realisierung solcher Vorhaben die Methode von Workshops vorzuschlagen, findet bei der Mehrzahl der befragten KI-Experten Zustimmung. So führt ein Experte seine Erfahrung an, dass besonders in KIGP häufig heterogene und diffuse Verständnisse von KI vorherrschten, die es vor dem Hintergrund potenziell hoher Zahlen betroffener Stakeholder zu diskutieren lohne:

*„Da gibt es eben verschiedene Sichtweisen auf die KI und verschiedene Anwendungsgebiete. Und man hat dann so ein breites Gebiet an Stakeholdern, dass da so ein Workshop wo die sich alle mal an einen Tisch setzen und ihre Seite darstellen und sich ihre Gedanken machen genau richtig. Weil man lernt ja auch dabei, wie die anderen über das Thema denken“ (Interview 4, Min 22:25).*

## 6.4 Schwächen

Die Schwächen beschreiben die Elemente der Methoden von CTA und VSD, die die KI-Experten in Bezug auf die Gestaltung von KI-Systemen kritisieren und als schwer umsetzbar oder schlicht nicht nachvollziehbar bewerten. Grundsätzlich haben die Experten im Vergleich mit den identifizierten Stärken nur wenige Schwächen identifiziert. Da viele der Experten in den Interviews die von ihnen identifizierten Stärken häufig ohne Anregung von meiner Seite überprüften, beziehen sich viele der nachfolgend dargestellten Schwächen direkt auf die zuvor dargelegten Stärken der Methoden.

### 6.4.1 Prozess- und Ergebnisorientierung

Ein zentraler Kritikpunkt an den CTA- und VSD-Methoden bezieht sich auf die Gewährleistung von konkreten Ergebnissen und Erkenntnissen, die sich für die Gestaltung von KI-Systemen als hilfreich und zielführend erweisen. Einige der KI-Experten bringen zum Ausdruck, dass es für viele Akteure in der Gestaltungspraxis notwendig sei, rasch die konkreten Vorteile des Einsatzes einer Methode zu erkennen und direkte Bezüge zu den Zielen eines Gestaltungsprojekts herzustellen. So fragt beispielsweise der KI-Experte in Interview 3:

*„Was ich mich auch tatsächlich gefragt habe ist: Gibt es Success Stories davon? Also gibt es [Beispiele] wo man das angewandt hat und dann gesagt hat „ja, jetzt haben wir bessere Fahrräder gebaut“ oder so etwas?“* (Interview 3, Min 12:16).

Er bezieht sich dabei auf die Anwendung der Methoden in Projekten, die nicht nur akademischen Charakter haben. Er gibt zu erkennen, dass solche Success Stories, also Beispiel- oder Pilotprojekte, die bewiesen haben, dass eine Methode funktioniert, zur Beurteilung sehr hilfreich sind. Zur Beurteilung möglicher Vorteile einer Methode werde häufig ein Blick auf die versprochenen Ergebnisse geworfen, wodurch die inhaltliche Argumentationsebene angesprochen ist. Diesbezüglich fiel den KI-Experten das Erkennen möglicher Ziele und Resultate der CTA- und VSD-Methoden schwer. In Verbindung damit steht häufig auch der Faktor Zeit im Mittelpunkt der Kritik an den Methoden, da alle KI-Experten den in der KI-Gestaltung vorhandenen Zeit- und Ergebnisdruck aus erster Hand kennen. Neben der inhaltlichen Rechtfertigung für den Einsatz einer Methode stehe häufig auch dieser mehr ökonomische Aspekt im Vordergrund. Da manche der Methoden keine konkreten Aussagen darüber treffen, wie viel Zeit bestimmte Handlungsschritte in Anspruch nehmen, um zielführende Ergebnisse hervorzubringen, fordern einige der KI-Experten konkretere Angaben, auch um den planerischen Aufgaben ihrer Berufsbilder gerecht werden zu können. Die genannten Schwächen der Methoden werden von den KI-Experten auf verschiedene Weise hervorgehoben und aus verschiedenen Perspektiven betrachtet. Auf die Frage, wie der beispielsweise der KI-Experte in Interview 9 die Anwendbarkeit der Methoden einschätzt, antwortet er:



*„Ich finde es akademisch sozusagen. Das hat bestimmt seinen Sinn, aber in der Realität [...] haben wir eigentlich nicht die Zeit, so exzessiv Methoden durchzugehen, weil jeder direkt schon Ergebnisse haben will. [Wir sollen] direkt mit Daten losgehen am ersten Tag [und] dann will man direkt auch irgendwelche Ergebnisse“ (Interview 9, Min 18:03).*

Hier wird also der Zielkonflikt aus Zeit- und Ergebnisdruck angesprochen, der in manchen KIGP bereits nach kurzer Zeit nach Projektbeginn vorherrscht. Die KI-Experten weisen stellenweise auch darauf hin, dass die Wahl der verwendeten Gestaltungsmethoden sich auf die benötigte Zeit für das Generieren von Gestaltungsergebnissen auswirkt. Über ein verzögertes Erreichen oder gar Ausbleiben von Ergebnissen muss gegenüber Kunden, Auftraggebern oder Geschäftspartnern häufig Rechenschaft abgelegt werden. Unter diesem Gesichtspunkt vermisst der KI-Experte in Interview 9 schlagende Argumente, die den Einsatz der Methoden rechtfertigen können. Der angesprochene akademische Charakter der Methoden trage nicht dazu bei, eine bereits heikle Situation zu entschärfen. Auf ähnliche Weise äußert auch der KI-Experte in Interview 3 Bedenken und kritisiert konkret die fehlenden Hilfestellungen, die die Anwendung der Methoden in Wirtschaftsprojekten aufzeigen oder erleichtern:

*„Letztendlich ist [KI-Gestaltung] Marktwirtschaft. [...] Wie kann das in so einer Welt überleben? Welches Incentive gibt es denn, das zu tun? [...] Wie bindet man das denn ein in kein Forschungsprojekt, sondern in ein wirtschaftliches Projekt? (Interview 3, Min 30:30).*

Der Zielkonflikt aus Zeit- und Ergebnisdruck wird auch in Interview 1 angesprochen. Hier erklärt der KI-Experte, dass bezüglich verfügbarer Zeit und dem notwendigen Erreichen von Gestaltungsfortschritt Kompromisslösungen gefunden werden müssen. Um diese zu herbeizuführen müssen Entscheidungen getroffen werden, die im Zweifel den Gestaltungsfortschritt vor Reflexions- oder Partizipationsmaßnahmen priorisieren:

*„The ideal is better to have everything at the earlier stage because later [the system] is more evolved. You have to find a trade-off because if you do everything beforehand, the project never starts. You have to find a balance [...], let's do it, let's start, release a beta, get feedback from the customers [...]. In industry there is this problem that you have deadlines“ (Interview 1, Min 29:48).*

Entsprechend werden die Methoden beider Ansätze bezüglich ihrer Unterstützung zur Herbeiführung von kompromissorientierten Lösungen zum Zeit- und Ergebnis-konflikt als wenig hilfreich aufgefasst, da dieser in beiden Ansätzen kaum thematisiert wird. Beide Ansätze widmen dem Thema Ressourcenmanagement grundsätzlich wenig Aufmerksamkeit, was sich entsprechend auf die vorgeschlagenen und entwickelten Methoden überträgt. Dies bemängeln einige der KI-Experten, da dies für sie in der Gestaltungspraxis ein zentrales Kriterium für die Auswahl und den Einsatz von Methoden darstellt. Daran anknüpfend wird auch der zielführende Beitrag der Durchführung von CTA-Workshops in Interview 1 in Frage gestellt:

*„What is more surprising out of the two for me would be the [CTA]-approach, because it actually involves these workshops. That is something interesting, but it also seems very expensive. So, I don't know if the industry is having that, I think not. The workshop would be sitting with a lawyer or sitting with customers and that's it“ (Interview 1, Min 18:26).*

*„I really like this VSD because you can keep these workshops out, that may be cumbersome or very difficult to organize“ (Interview 1, Min 28:16).*

Eine weitere Schwäche der CTA-Methoden identifiziert der KI-Experte aus Interview 8 unter dem Gesichtspunkt, dass den Methoden Maßnahmen fehlen würden, die sich auf die Verwertung der mit den Methoden erarbeiteten Ergebnisse beziehen. Dem Experten fehlen hier konkrete Angaben, wie die Ergebnisse allgemein, speziell jedoch in KIGP verwendet werden:

*„Hier [in CTA] fehlt also wirklich dieser Feedback-Loop. Man hat jetzt den Workshop durchgeführt, man hat das bewertet, man hat vielleicht auch Hypothesen abgeleitet [...]. Ich glaube da fehlt aber [die Möglichkeit] zum Experimentieren, Ausprobieren, Pilotieren und dann wieder so eine Nachbetrachtung zu machen (Interview 8, Min 28:31).*

Er schlägt die Ergänzung von Feedback-Loops vor, in deren Rahmen aus den entwickelten Szenarien Gestaltungsmaßnahmen abgeleitet werden und deren Umsetzung in Pilotphasen erprobt und getestet werden. Die Ergebnisse solcher Pilotphasen können anschließend wiederum mit den Inhalten der zuvor entwickelten Szenarien verglichen werden, um anhand der tatsächlich eingetretenen Entwicklungen weitere Rückschlüsse und Entscheidungen für das zukünftige Gestaltungsvorhaben abzuleiten. Der KI-Experte gibt darüber hinaus zu verstehen, dass die Entwicklung von Szenarien in wirtschaftlichen Projektkontexten selten aus Selbstzweck geschehe, sondern mit einer Szenarienenentwicklung immer ein erkennbarer Nutzen und Mehrwert für das Projekt verbunden sein müsse.

Auch in Bezug auf die VSD-Methoden werden von den KI-Experten Schwächen benannt und aufgezeigt. So beschreibt der KI-Experte in Interview 8, dass ihm die reine Identifikation potenzieller Risiken eines KI-Systems im Rahmen einer Stakeholderanalyse nicht weit genug geht. Ihm fehlt an dieser Stelle der nächste und entscheidende Schritt, anhand der identifizierten Risiken konkrete Gestaltungsmaßnahmen zu formulieren und umzusetzen:

*„Ich glaube hier braucht es auch eine Art von Maßnahmenkatalog. Also nur die Risiken zu identifizieren reicht da glaube ich nicht, sondern man müsste dann schon ableiten – wie geht man mit den Risiken um? Akzeptiert man die, oder gibt es eine Maßnahme um die zu mildern oder zu vermeiden? [...] Sonst fällt das auf dem Weg zum Projektziel auch irgendwo runter“ (Interview 8, Min 37:07).*

Auf ähnliche Weise bemängelt der KI-Experte in Interview 5 die CTA-Methode. Er spricht davon, dass die Methode einerseits dazu in der Lage ist, das Bewusstsein aller Beteiligten auf größere Fragen und Zusammenhänge auszurichten, dass jedoch am Ende der weitere Schritt fehle, aus den CTA-Workshops konkrete Maßnahmen für einen Gestaltungsprozess abzuleiten. Er bezeichnet diese Tätigkeit als das Erbringen einer Abbildungsleistung auf die technologische Domäne:

*„Vor allem die [CTA-]Methode, das bettet das Ganze schön ein und man hat danach bestimmt ein größeres Bewusstsein und einen breiteren Horizont, aber es ist halt wenig pragmatisch. Also da muss tatsächlich noch eine Abbildungsleistung erfolgen auf die technologische Domäne. Und das geschieht hier eher nicht“* (Interview 5, Min 28:52).

Am deutlichsten wird der Mangel der Ergebnisorientierung der Methoden beider Ansätze jedoch in Interview 9 kritisiert. Der KI-Experte erläutert dazu (sehr leidenschaftlich) die Denkweisen und den Berufsethos von vielen Technikgestaltern, die sich durch einen starken Fokus auf das Erschaffen und Gestalten tatsächlicher technischer Artefakte auszeichnen. Gestalter wollten Dinge „bauen und zeigen“ und bräuchten für die Umsetzung dieser Bestrebungen, die häufig auch den Anspruch an die Berufsgruppen von Auftrag- und Arbeitgebern darstellen, ergebnisorientierte Methoden an die Hand:

*„Ich glaube die Methodik versteht sozusagen nicht, wie stark dieser Fokus tatsächlich auf der technischen Sache ist, immer zu irgendwelchen Ergebnissen zu kommen, die man dann wirklich besprechen kann. Dass das sozusagen so wichtig ist“* (Interview 9, Min 35:00).

Die Ergebnisorientierung und das Unterstützen der damit verbundenen Einstellungen und Bedürfnisse sei in den Methoden beider Ansätze nicht vorhanden, weshalb der KI-Experte in der Anwendung der Methoden in KIGP kaum einen Mehrwert erkennen kann.

## 6.4.2 Befähigung zur Reflexion

Auch wenn in der Mehrzahl der Experteninterviews die Befähigung zur Reflexion als eine Stärke der Methoden bezeichnet wurde, äußern manche der KI-Experten unter demselben Gesichtspunkt Zweifel hinsichtlich der Nutzbarkeit und des Mehrwerts dieses Effekts in bzw. für KIGP. Die Gründe für diese Zweifel sind verschiedentlich verortet und werden von den KI-Experten in den meisten Fällen durch Verweise auf den Gestaltungsalltag in KI-Projekten erklärt. So beschreibt beispielsweise der KI-Experte in Interview 3, dass der Gestaltungsalltag häufig erfordere, sich mit konkreten Problemen und Herausforderungen auseinanderzusetzen und dass es deshalb schwierig sei, die großen Zusammenhänge immer im Blick zu behalten. Man müsse Prioritäten setzen, und im Zweifel habe die technische Funktionalität oder ein Problem-Anruf vom Vorgesetzten oder Kunden höhere Priorität wie z.B. die Frage, ob ein bestimmtes Feature möglicherweise negative gesellschaftliche Auswirkungen zur Folge haben könnte:

*„[Es ist] halt bisschen schwierig [...]. Im täglichen Doing macht man das ja nicht. [...] Man arbeitet ja immer an Details, man arbeitet ja nie am großen Ganzen“*  
(Interview 3, Min. 18:13).

Er spricht damit einen Aspekt an, der auch in anderen Interviews und in anderen Zusammenhängen immer wieder zur Sprache kommt – die Frage nach den Verantwortungsreichen in KIGP. Es geht dabei um die – auch in HCAI – zentrale Frage, inwieweit es angemessen ist, Gestalter für die Adressierung und Berücksichtigung sozialer und ethischer Aspekte verantwortlich zu machen und (siehe Kap. 6.7.5). Auf diesen Aspekt geht auch der KI-Experte in Interview 5 ein, er weist bei der Verantwortungsfrage jedoch auf den Gesichtspunkt hin, dass insbesondere die VSD-Methode der Value dams and flows den Gestaltern und Entscheidungsträgern bereits einen großen Vertrauensvorschuss entgegenbringen würde, da sie damit beauftragt würden, die Nutzen und Risiken des von ihnen selbst gestalteten Systems zu identifizieren. Diesen Vertrauensvorschuss hält der Experte für fragwürdig:

*„Da muss man schon ein großes Vertrauen in das Projektteam haben, das wird dann wahrscheinlich nicht der junge Abiturient sein. Da wird halt eine gewisse Autorität vorausgesetzt im Projektteam, die das Projektteam möglicherweise nicht hat“* (Interview 5, Min. 30:00).

Er spricht hier demnach zwei Schwächen der Methode an – zum einen kritisiert er, dass keine genaueren Angaben dazu gemacht werden, wer die Nutzen und Risiken identifizieren soll und dass für die Durchführung dieser Aufgabe von der Methode kein Mindestmaß an Berufs- und Lebenserfahrung vorausgesetzt wird. Darüber hinaus deutet er die Problematik an, dass Projektteams einem gewissen Bias unterliegen, wenn sie ein technisches Artefakt beurteilen sollen, dass sie größtenteils selbst gestaltet haben. Dadurch sind sie mit hoher Wahrscheinlichkeit unfähig, die Biases zu erkennen und für potenziell kritisch zu befinden, die aufgrund ihrer eigenen Werte und Überzeugungen durch ihre Beteiligung an der Gestaltung in das System miteingeflossen sind. Er verdeutlicht dies etwas später im Interview erneut mit der Aussage:

*„Das Projektteam ist relativ allmächtig“* (Interview 5, Min 35:24).

Er deutet damit an, dass die Identifikation potenzieller Nutzen und Risiken eines technischen Systems nicht durch die Gruppen und Individuen erfolgen sollte, die maßgeblich für seine Gestaltung verantwortlich sind.

Auch in Interview 4 kommt das Bestreben der Methoden, eine ganzheitliche Perspektive auf das Gestaltungsgeschehen zu etablieren zur Sprache. Allerdings identifiziert der KI-Experte in diesem Interview an dieser Bestrebung eine starke Einseitigkeit, da ihm die nicht-intendierten Nutzen von KI-Systemen in den Methoden beider Ansätze zu stark im Vordergrund stehen. Unter diesem Gesichtspunkt erklärt er auch, dass in seiner Wahrnehmung der Ansatz VSD auch Teil der CTA-Methodik sein könnte:

*„Das [VSD] ist für mich eine Teilmenge vom [CTA]. Ich finde es ein wenig unausführlich. Da steht zwar „nicht-intendierter Nutzen“ [...] aber ich finde etwas ganzheitlicheres, eine SWOT-Analyse oder so besser. Mir ist das [VSD] einfach zu dünn. [...] Es fühlt sich für mich so an, als wenn da ganz viel fehlt. Es ist nicht so ganzheitlich drauf geschaut. Warum nur nicht-intendierter Nutzen, warum nicht auch der intendierte Nutzen? Warum ist der außen vor?“ (Interview 4, Min 26:20).*

Wie aus diesen identifizierten Schwächen deutlich wird, verwenden die KI-Experten ihre in der Gestaltungspraxis erlebten Erfahrungen dazu, die aus ihrer Sicht zu ambitionierten oder illusorischen Elemente und Annahmen der Methoden zu benennen und zu entkräften.

### 6.4.3 Partizipation von Stakeholdern

Auch zum Partizipationsaspekt identifizieren die KI-Experten vereinzelte Schwächen. Wie bereits zum Aspekt des Vertrauensvorschlusses der VSD-Methode findet der KI-Experte in Interview 5 auch in Bezug auf die CTA-Methoden deutliche Worte und kritisiert den zielführenden Beitrag, den CTA-Workshops zu KI-Gestaltungsprozessen zu leisten in der Lage sind:

*„Der [CTA]-Ansatz ist, und das klingt jetzt sehr abschätzig, aber es ist sehr laborlastig. Also no offense, aber ich kenn das ja. Ich war schon bei Startups, bei einem Konzern und es wird halt relativ viel zerredet. Und wenn die Runden relativ groß sind, dann hast du oftmals ein, zwei, drei dominante Stimmen und viele werden dann gar nicht gehört, weil sie sich nicht trauen, schüchtern sind, Angst haben. So große moderierte Workshops begünstigen halt oftmals die Leute, die besonders laut sind. Ich find das auf [...] den ersten Blick schön, aber das ist auch Macht versteckt“ (Interview 5, Min 35:24).*

Er spricht hier den Aspekt der Machtfrage an, und zweifelt daran, dass es in den CTA-Workshops gelingen kann, mit möglicherweise vorhandenen Machtstrukturen (z.B. durch bestehende Hierarchien innerhalb einer Organisation) zielführend umzugehen und ihren negativen Tendenzen durch Mitteln der Moderation Einhalt zu gebieten. Auch benennt er die Schwäche der Methode, keine Maßnahmen vorzusehen, um möglicherweise angestrebten Verzögerungs- oder Verschleierungsstrategien von Workshopteilnehmern entgegenzuwirken. Während diese Schwäche sicherlich auf das Workshopformat im Allgemeinen übertragen werden kann, ist die Benennung der Machtstrukturen im Zusammenhang mit CTA bemerkenswert, da Machtstrukturen und Machtgefälle in sozialen Gruppen, gesellschaftlichen Institutionen und auch in Gestaltungsprozessen (Impactors und Impactees) in der CTA-Literatur intensiv diskutiert werden. Im Zusammenhang der in CTA vorgeschlagenen Methodik wird dieser Aspekt hingegen nur wenig beleuchtet, wodurch es sich dabei um einen weiteren blinden Fleck in CTA handelt. Eine weitere Schwäche des CTA-Workshopformats identifiziert der KI-Experte in

Interview 6, wo er anzweifelt, dass die Workshopteilnehmer in der Lage seien, sich gedanklich angemessen zielführend in die Zukunft zu denken:

*„Das ist aber ziemlich weit in der Zukunft, die können sich das nicht vorstellen“*  
(Interview 6, Min 16:42).

Entsprechend sieht er eine Schwäche der CTA-Methodik darin, ihr inhaltliches Ergebnis und ihren strategischen Beitrag zu Gestaltungsprozessen stark von der Imaginationsfähigkeit der Workshopteilnehmer abhängig zu machen. Für ihn handelt es sich dabei um eine deutliche Schwäche, da im Vorfeld von den Anwendern der Methode nur vage antizipiert werden kann, welche Zukunftsszenarien die Workshopteilnehmer ersinnen werden und für wie wahrscheinlich ihr Eintreten sich letztlich für den Gestaltungsprozess erweist. Zur VSD-Methode des Value-oriented Prototyping ergänzt der Experte später im Interview eine weitere Schwäche, die eine ähnliche Kritik aufweist:

*„Die Teilnehmer die können ja auch nicht immer alles ausdrücken, [...] d.h. die Transformation von dem Prototypen über die Analyse zu den Stakeholderwerten, da ist definitiv viel Interpretationsspielraum. Sehr viel Interpretationsspielraum“* (Interview 6, Min 34:01).

Auch hier spricht der Experte zunächst die Limitierung des Partizipationsformats an, die darin besteht, dass ihre Ergebnisqualität stark an die kognitiven und verbalen Fähigkeiten der Stakeholder gebunden ist. Sind die Stakeholder nicht in der Lage, ihre Gedanken auf eine Weise auszudrücken oder zu demonstrieren, die eine Ableitung von relevanten Aspekten und Faktoren für einen Gestaltungsprozess ermöglichen, wurden für die Durchführung der Methoden Zeit aufgewendet, ohne handhabbare Ergebnisse zu produzieren. Daran anknüpfend sieht der KI-Experte eine weitere Schwäche der Methode in der Interpretations- und Transferleistung, die das Projektteam in der Methode erbringen soll. Das Projektteam soll anhand der von Stakeholdern getätigten Aussagen und gestalteten Prototypen die für Stakeholder bedeutsamen Werte ableiten, insbesondere jene Werte, die den Stakeholdern selbst nur in geringem Umfang bewusst sind. Hier kritisiert der Experte den großen Interpretationsspielraum der dem Projektteam zur Verfügung steht und bemängelt entsprechend die Abwesenheit von Maßnahmen zur Beschränkung möglicherweise vorhandener Biases innerhalb des Projektteams. Ähnlich wie der KI-Experte in Interview 5 kritisiert er damit die dem Projektteam zugewiesene hohe Verantwortung, bedeutsame und im Zweifel sensible Entscheidungen zu Wertefragen treffen zu können und zu müssen, ohne diese an spezifischen Kriterien auszurichten zu müssen.

## 6.5 Chancen

Die Chancen beschreiben die Elemente der Methoden von CTA und VSD, die die KI-Experten in Bezug auf die Gestaltung von KI-Systemen als im positiven Sinne beitragsfähig beurteilen. Sie beschreiben verschiedene Potenziale, die die Methoden in ihrer in den Interviews dargestellten Form nicht oder nur wenig thematisieren und adressieren, die jedoch nach Einschätzung der Experten zielführende und wünschenswerte Effekte auf Prozesse der KI-Gestaltung haben können. Da die Mehrzahl der Experten durch die Interviewstudie erstmalig in Kontakt mit den CTA- und VSD-Methoden gekommen sind, sind die Aussagen zu Chancen und Risiken im Vergleich mit identifizierten Stärken und Schwächen vielfältiger und differenzierter.

### 6.5.1 Prozess- und Ergebnisorientierung

In Bezug auf die Eignung für Prozesse der KI-Gestaltung sprechen einige der KI-Experten den Methoden eine grundsätzliche Chance zielführender Beiträge aus. Die Begründungen und Detailebenen, auf die sie für ihre Einschätzungen zurückgreifen variieren, doch sind die diesbezüglichen Aussagen eindeutig. So sprechen sich die KI-Experten aus den Interviews 3 und 7 recht grundsätzlich dafür aus, dass die Anwendung der CTA- und VSD-Methoden in KI-Projekten vernünftig und zielführend sei:

*„Das macht schon Sinn und es ist vernünftig, dass man das macht. Ich glaube das wird in der Industrie nicht gemacht [...], aber man sollte es sicher tun“* (Interview 3, Min 11:50).

*„Ich muss sagen ich finde [KI] eigentlich einen perfekten Use Case. Von den Ansätzen und den Ideen [...] finde ich passt es sehr sehr gut in dem Rahmen“* (Interview 7, Min 40:01).

Der KI-Experte aus Interview 3 bezieht sich in seiner Chancenzuschreibung sehr auf die Befähigung zur Reflexion der Methoden, während für den Experten aus Interview 7 primär die Möglichkeiten zur Stakeholderpartizipation ausschlaggebend sind, die die Methoden vorsehen. Für ihn stehen in der KI-Gestaltung besonders die übergreifenden gesellschaftlichen Themen- und Fragestellungen im Vordergrund, weshalb er insbesondere die KI-Gestaltung als sehr wichtiges Gestaltungsfeld sieht, in dem derartige Fragestellungen mitgedacht und besonders berücksichtigt werden sollten:

*„Wenn man über Automatisierung, Roboter in Produktionsanlagen, Autonomes Fahren usw. [spricht], was da alles eigentlich auf der Kippe steht, v.a. auch im sozialen Umfeld – was macht man dann mit den ganzen Taxifahrern, mit den LKW-Fahrern, wenn alles automatisiert umgesetzt wird? Und ich glaube da hat genau das KI-Umfeld die richtigen Use Cases um eben solche Konzepte dann auch anzuwenden“* (Interview 7, Min 40:38).

Der KI-Experte aus Interview 4 teilt diese Einschätzung und bezieht sich in seiner Argumentation primär darauf, dass insbesondere die CTA-Methoden darauf setzen, ein möglichst breites Feld an Technologiefeldakteuren und Workshopteilnehmern zu gewinnen und deren Einsichten und Perspektiven zu ermitteln. Diese Tätigkeit, so der Experte, erfahre insbesondere bei der Gestaltung von KI-Systemen eine sehr hohe Bedeutung, da die Technologie auf sehr viele Lebensbereiche angewendet werden könne, und man deshalb möglichst viele Stakeholderperspektiven berücksichtigen müsse:

*„Es gibt ja tausend Dinge für die man KI einsetzen kann, wo man einfach ganz verschiedene Stakeholder hat. Und für so ein Feld, das wirklich keine kleine, sondern eine große Gruppe an Stakeholdern hat, aus allen Lebensbereichen, aus allen Fachbereichen, da finde ich so ein Vorgehen genau richtig – man holt sich eben sämtliche verschiedenen Sichtweisen“* (Interview 4, Min 21:40).

In Interview 5 sieht der KI-Experte ebenfalls eine Chance für eine Beitragsfähigkeit der VSD-Methoden zu KIGP, er knüpft seine Einschätzung jedoch an zwei Bedingungen. Zum einen setzt er voraus, dass in Schritt 1 der Value dams and flows-Methode, die Nutzen und Risiken eines KI-Systems objektiv ermittelt werden, zum anderen sei zu bedenken, dass es sich bei Fragebögen um psychometrische Instrumente handle, die seiner Erfahrung gerne mit Messfehlern oder Verzerrungen in Verbindung stünden:

*„Wenn Nutzen und Risiken objektiv etabliert sind, dann [halte] ich die Methode der Fragebögen für ausgezeichnet. Als Statistiker kann ich noch anmerken: Es ist ja auch ein psychometrisches Mittel, das müsste man dann auch entsprechend noch eichen und sauber einstellen. [...] Dass auch die Fragebögen einer gewissen Wissenschaftlichkeit genügen müssten, das wäre wünschenswert“* (Interview 5, Min 31:17).

Mehr gegen Ende des Interviews 9 gesteht auch der ansonsten eher kritische KI-Experte den Methoden die Chance zu, in KI-Gestaltungsprozessen insbesondere dazu beizutragen, systematisch Folgen abzuschätzen. Dies sei eine Tätigkeit, die in der Gestaltungspraxis deutlich zu wenig Aufmerksamkeit erfahre, die für Gestalter und Entscheidungsträger jedoch häufig viele wertvolle Informationen und Impulse bereithalte:

*„Ja, ich glaube es kommt auf jeden Fall auch zu kurz, dass wir uns nicht Gedanken darüber machen, [...] Folgen abzuschätzen. Wir machen das nicht so systematisch [...]. Das könnte man bestimmt nochmal mitnehmen. Ich finde das ist eine gute Idee [...]. Aber ich glaube das wäre total sinnvoll, weil wir uns nicht alle Eventualitäten überlegen, die dann vorkommen“* (Interview 9, Min 30:50).



## 6.5.2 Wissenstransfer und Werteorientierung

In Interview 1 wird den Methoden beider Ansätze vom KI-Experten das Potenzial zugeschrieben, gut geeignet zu sein, um einen Wissenstransfer zwischen verschiedenen Fachbereichen, sowie zwischen verschiedenen gesellschaftlichen Institutionen zu ermöglichen. Der Experte erläutert dies in Referenz zum gedanklichen Konstrukt eines virtuous cycle (z.B. Zahra, 2008), der im Englischen eine Kette von wünschenswerten, sich selbst kontinuierlich verstärkenden Ereignissen beschreibt (Gegenteil eines Teufelskreislaufs):

*„The objective of every AI-system is not just performance, but also fairness etc. I think [these approaches] are good for educational purposes, for really having the systems doing what we want them to do. This has to start in the universities then going back to society, from society to the law, from the law to the industry. And if possible, people that come to the industry come from a society that already has the proper mindset and they don't lose it in the industry. So it would be like a virtuous cycle, taking these things into account“ (Interview 1, Min. 25:15).*

Er verbindet mit den Methoden also die Chance darauf, in Gesellschaften langfristig tiefgreifende und positive Veränderungen und Werteeinstellungen herbeizuführen, setzt dabei jedoch voraus, dass nachfolgende Generationen von KI-Gestaltern, Entscheidungsträgern und anderen relevanten Gruppen mit den Methoden an Universitäten in Kontakt kommen und dort ihre Anwendung erlernen. Später im Interview ergänzt er seine Argumentation um einen weiteren Aspekt, denn für ihn bergen die Methoden auch das Potenzial die Funktion von „intermediate gates“ auszuüben. Damit bezeichnet er auf metaphorischer Ebene eine Zwischentorfunktion, um eine Verbindung zwischen Gestaltungstheorie und Gestaltungspraxis herzustellen, um einen Wissens- und Erfahrungstransfer zu ermöglichen. Besonders unter diesem Gesichtspunkt bergen die Methoden für ihn ein wertvolles Potenzial, dass sie bislang wenig ausschöpfen:

*„I think it's very important to really think about it. It's not only that we need philosophers thinking about the more purely theoretical things – what is more ethical [in AI] – but we always also need an intermediate gate between the pure concepts and the industry. I think approaches like this can communicate both of them“ (Interview 1, Min 32:20).*

In Interview 2 wird den Ansätzen das Potenzial zugeschrieben, zur Aufklärung über die Funktionsweise von KI-Systemen beizutragen und dadurch wiederum die Steigerung der Akzeptanz von KI-Systemen zu fördern. Der KI-Experte verwendet in seiner Begründung dieser Einschätzung Erfahrungswerte aus der Gestaltungspraxis, in der er selbst häufig derartige Aufklärungsarbeit leisten muss, und immer wieder erstaunt darüber ist, wie (aus seiner Sicht) abwegige Vorbehalte gegenüber KI-Systemen geäußert werden:

*„Wir haben auch immer wieder Probleme mit der Akzeptanz in der Gesellschaft. [...] Was wir immer wieder sehen ist, dass die Mitarbeiter eine hohe Skepsis ha-*



*ben, es ist halt nicht so einfach nachzuvollziehen wie ein neuronales Netz funktioniert“ (Interview 2, Min 14:10).*

Er gesteht ein, dass die in KI-Systemen verwendeten Verfahren häufig schwer nachzuvollziehen seien, besonders, wenn keine technische Ausbildung absolviert wurde. Aufschlussreich ist hingegen seine Assoziation, dass die CTA- und VSD-Methoden dazu verwendet werden könnten, diese technische Komplexität zu reduzieren, transparent zu machen, und verständlich zu kommunizieren. Dies würde einen positiven Nebeneffekt darstellen, da weder in CTA noch VSD sich eine derartige Funktion explizit zum Ziel gesetzt wird. Das Transparentmachen der Funktionsweise von KI-Systemen könnte aus Sicht des KI-Experten Gegenstand von beispielsweise CTA-Workshops sein, um Aufklärungsarbeit zu leisten und dazu beizutragen, die möglicherweise falsch informierte oder fehlgeleitete Skepsis bei Stakeholdern zu verdeutlichen:

*„Niemand möchte fehlerhafte Teile ausliefern, die Kosten sind sofort brutal [...], wenn man chargenweise wegwerfen soll. Und dann zu sagen, okay wir haben da jetzt so eine KI, die niemand richtig versteht bei uns, die das macht, von einem Zulieferer eingekauft, dann ist das natürlich erst einmal eine Skepsis, die man überwinden muss. Und das fällt uns schon schwer“ (Interview 2, Min 15:12).*

Er deutet in seiner Aussage einen Zweck der Methoden auch dahingehend um, sie als strategisches Werkzeug einzusetzen, durch deren aufklärerische Funktion sie auch als Werkzeug zur Einkaufsüberzeugung genutzt werden können. Für ihn steht damit nicht nur der technische, sondern auch der ökonomische Mehrwert im Vordergrund, die er den Methoden zu leisten zuerkennt.

### 6.5.3 Kompatibilität und Variabilität

Die KI-Experten bringen in den Interviews auch verschiedentlich zum Ausdruck, dass sich die CTA- und VSD-Methoden auch im Verbund einsetzen ließen. Dies bringt beispielsweise der KI-Experte in Interview 5 zum Ausdruck, wenn er sagt:

*„Natürlich würden sich auch beide Methoden kombinieren lassen“ (Interview 5, Min 23:40).*

Er bezieht sich dabei vor allem auf die Eignung beider Methoden, für Menschen relevante Werte zu ermitteln und diese im Gestaltungsgeschehen zu berücksichtigen. Auf ähnliche Weise äußert sich der KI-Experte in Interview 6. Er greift in der Argumentation seiner Einschätzung auf seine Erfahrung im Feld der Gestaltung hybrider Systeme (Software & KI) zurück, und erläutert, dass eine hohe Flexibilität auf technischer wie auch auf methodischer Ebene in Kontexten der KI-Gestaltung von großem Vorteil sein kann. Auch er identifiziert in den Methoden von CTA und VSD die Chance, sich gut zu ergänzen und an spezifische Projektgegebenheiten oder angestrebte technische Anwendungsfälle anpassen zu lassen:

*„Beide kann man verwenden, wahrscheinlich in einer hybriden Art und Weise, je nachdem für welche Fälle [ein System] eingesetzt wird. [...] In dem einen Fall ist die eine Methode, in dem anderen die andere und in manchen Fällen sind beide in Kombination eine Lösung“ (Interview 6, Min 25:24).*

Noch konkreter kann er sich vorstellen, die CTA-Workshops mit der Methode des Value-oriented Prototypings zu kombinieren:

*„Das sind so Aspekte von der einen Methode, die ich adaptiv einsetzen kann und sagen kann, ich kombinier die jetzt. [...] Da sagt man, okay wir haben jetzt die Prototypen und verknüpfen die mit Elementen aus dem [CTA-]Workshop. Ich denke das wäre durchaus sinnvoller, als die getrennt zu betrachten“ (Interview 6, Min 27:15).*

Er argumentiert explizit dafür, beide Methoden in Verbindung einzusetzen und nicht getrennt zu betrachten. Somit wird an dieser Stelle deutlich, dass der Anspruch beider Ansätze erfüllt wird, methodisch auf möglichst viele Gestaltungsziele anwendbar zu sein und zu zusätzlichen Gestaltungsphilosophien kompatibel zu sein.

#### 6.5.4 Befähigung zur Reflexion und Folgenabschätzung

Zum Aspekt der Befähigung zur Reflexion haben die KI-Experten geteilte Einschätzungen geäußert. Einige der Bestrebungen und Vorschläge der Methoden identifizieren sie explizit als Stärken, andere hingegen als eindeutige Schwächen. Darüber hinaus identifizieren viele KI-Experten in den Methoden ein Potenzial zur Befähigung zur Reflexion, die sich in sehr unterschiedlichen Facetten äußert. Hinsichtlich der Befähigung zur Reflexion durch Szenarien und Szenarienentwicklung (CTA) rekurriert der KI-Experte in Interview 4 auf seine Erfahrungswerte im Umgang mit Szenarien in Projekten der Vergangenheit:

*„[Dort] haben wir uns mit Experten aus ganz verschiedenen Ecken und Bereichen [...] die verschiedenen Szenarien vorgenommen und überlegt – Was würde das bedeuten? Was würde das für Produkte am Markt erzeugen? Was würde es für Risiken [...], Probleme und Möglichkeiten erzeugen? Das war super interessant und hat super viel Spaß gemacht, und gerade bei KI kann ich mir das auch sehr gut vorstellen“ (Interview 4, Min 13:06).*

Der KI-Experte spricht der Entwicklung und Diskussion von Zukunftsszenarien das Potenzial zu, zielführende Beiträge zu KI-Gestaltungsprozessen zu leisten, unter der Voraussetzung, dass diese in einem fachlich heterogen besetzten Plenum diskutiert werden können. Auch wird bei dieser Aussage deutlich, dass der KI-Experte zusätzlich zu produktiven auch freudvolle Erfahrungen mit der Arbeit mit Szenarien verbindet, und dadurch eine weitere Chance dieser Methode zum Ausdruck bringt. Eine weitere Chance, die in Interview 6 mit der Befähigung zur Reflexion assoziiert wird, bezieht sich auf den Aspekt, gesellschaftliche Bedenken zu einer Technologie zu ermitteln und diese systematisch in Gestaltungsprozessen zu berücksichtigen:

*„Vorausgesetzt wir haben [Ängste] festgestellt, dann macht es schon Sinn [...], dass man sagt okay, [...] welche Auswirkungen hat jetzt diese Technologie auf die Gesellschaft? Da macht es definitiv Sinn, in Bürgerforen oder auch in Diskussionsrunden [...], dass die Teilnehmer die Bedenken äußern. [So] können die Bedenken gezielt adressiert werden“ (Interview 6, Min 15:02).*

Der Experte macht jedoch zur Bedingung, dass im Verhalten von Stakeholdern Vorbehalte gegenüber einer Technologie erkennbar sein müssen, um dieser Tätigkeit nachzugehen. Er positioniert sich vorsichtig optimistisch, dass die Methoden einen Beitrag leisten können, gesellschaftlich relevante Aspekte in den Vordergrund zu rücken, da er um die hohe Komplexität weiß, die derartige Auseinandersetzungen typischerweise mit sich bringen. Er weiß jedoch auch um die Mehrwerte, die ein interdisziplinärer Austausch zu gesellschaftlichen Auswirkungen von TG hervorbringen kann:

*„Das ist halt eine ziemlich komplexe Aufgabe und ich denke da hilft es oder ist wahrscheinlich wertvoll, das interdisziplinäre Unabhängige“ (Interview 6, Min 17:10).*

Er befindet dafür insbesondere die in CTA in Anspruch genommene Unabhängigkeit von in der TG aktiven Gruppen als zielführend. Eine weitere Chance der CTA- und VSD-Methoden sieht der KI-Experte darin, die Annahmen von Stakeholdern zu einer Technologie freizulegen, und um über auf unvollständigen oder fehlerhaften Informationen basierende Annahmen aufzuklären. Er verwendet zur Erläuterung seiner Argumentation das Beispiel, dass viele Menschen KI-Systeme mit der Annahme assoziieren, dass die Systeme immer zu einem eigenständigen Lernen in der Lage sind. Dies führe zu Vorbehalten und Ängsten, die häufig völlig unbegründet seien:

*„Wenn wir unsere Systeme präsentieren, dann hören wir immer: „das ist ja KI, das lernt ja von selbst.“ Dann [stellen wir] die Frage: „Woher haben Sie die Information?“ Die Menschen sind so gebiased mit Informationen [...]. Die News werden ja heute geschrieben um Klicks zu generieren, das ist fürchterlich [...]. Das heißt, da ist so viel Aversion schon in den Markt reingeflossen, [...] ein Priming das dazu führt, dass die Leute direkt schon eine Gegenposition einnehmen zu einer Technologie“ (Interview 6, Min 37:02).*

Auch er beschreibt, dass es seine beruflichen Tätigkeiten häufig mit sich bringen, Aufklärungsarbeit über die KI-Technologie mit ihren Möglichkeiten und Limitierungen zu leisten. Aufgrund dieser Erfahrung weiß er, dass Ängste und Vorbehalte von Stakeholdern sehr ernst genommen werden müssen, besonders wenn ersichtlich wird, dass diese Ängste und Vorbehalte auf unvollständigen oder gar fehlerhaften Informationen beruhen:

*„Und das haben wir gemerkt, man muss tatsächlich offen mit [Fachanwendern] reden und auch offen entwickeln und ihnen die Ängste nehmen“ (Interview 6, Min 39:17).*

Da er um diese Umstände weiß, begrüßt er die in den Methoden vorhandenen Möglichkeiten, Ängste zu identifizieren und ihre Ursprünge und Hintergründe zu beleuchten. Daran anknüpfend sprechen die KI-Experten in Interview 1 und 8 den Aspekt von Unsicherheiten an, bei deren Identifikation und Beseitigung die Methoden von CTA und VSD unterstützen können:

*„I would only do [CTA] if I have a lot of doubt [...]. If there are some dark areas we are really not sure about, then I would go through the workshop part“* (Interview 1, Min 28:48).

*„Jede AI-Initiative die wir haben, in denen wir Geschäftsprozesse ändern wollen, in denen wir unsere Mitarbeiter aber auch Kunden mit KI konfrontieren wollen, haben einen Projektcharakter. Da ist Unsicherheit von Anfang an da – was möchte ich erreichen damit? Welche Kennzahl möchte ich verbessern? [...] Funktioniert das mit den Daten die wir haben? [...] Da ist ganz viel Unsicherheit da, und deswegen braucht es einen Projektmanagementansatz“* (Interview 8, Min 21:20).

Der KI-Experte in Interview 8 nimmt hier Bezug auf einen Projektmanagementansatz, dessen Bestandteile nach seiner Einschätzung die CTA- und VSD-Methoden darstellen könnten. Somit sieht er in beiden Methoden auch das Potenzial, dass sie neben den gestalterisch-inhaltlichen Aspekten auch zu den Managementaspekten von KIGP beitragen können. Dies lässt darauf schließen, dass für den Experten die in den Methoden verankerte strukturierte Vorgehensweise weitgehend nachvollziehbar wirkt. Insbesondere aufgrund seiner Erfahrungswerte im Umgang mit Stakeholderanalysen schätzt er die Chance von VSD recht hoch ein, wertvolle Erkenntnisse für KIGP beizutragen. Gestärkt wird seine Einschätzung vom KI-Experten in Interview 9. Hier sieht der KI-Experte eine Chance der Methoden darin, Räume für ruhige und reflektierte Gespräche zu schaffen. Er gesteht jedoch auch ein, dass dies nach seiner Erfahrung sehr viel Kraft erfordere und grundsätzlich nur dann erfolgen könne, wenn entsprechende Budgets freigegeben worden seien:

*„Es ist super schwierig dann nochmal ein wenig Ruhe reinzubringen [im Sinne von] „komm wir sprechen nochmal ganz in Ruhe über Dinge, versuchen uns vielleicht nochmal zusammenzusetzen.“ Das bedarf dann sehr viel Kraft“* (Interview 9, Min 18:48).

Der Experte erkennt in den Methoden die Chance, Gestaltungsprozesse zu entschleunigen, um tiefgreifende und deshalb relevante Aspekte mit Mitarbeitern und Auftraggebern gemeinsam zu diskutieren. Der Kraftaufwand, den der Experte anspricht, bezieht sich auf die häufig an Deadlines orientierten Projektpläne, die häufig nur wenig Spielräume für Reflexion und Diskussion zulassen. Im Bezug zu den CTA-Workshops erkennt er die Chance, durch Stimmen „von außen“ wertvolle Einsichten zu generieren. Er weist darauf hin, dass dies „wichtig wäre“, es aber nach seiner Einschätzung nicht gemacht werde (Interview 9, Min 21:40). Der Experte ergänzt seine Einschätzung etwas später im Interview,

und hebt dabei besonders das Potenzial der Methoden hervor, vorausschauend wertvolle Einsichten zu ermitteln:



*„Was in einem KI-Projekt charakteristisch ist, ist, dass es sehr iterativ ist, dass wir am Anfang eben nicht diese Gespräche hatten mit z.B. Vertriebsmitarbeitern. Wenn wir sowas ganz am Anfang gemacht hätten, wäre es uns nicht aufgefallen, sondern das bemerken wir erst im Laufe der Zeit“* (Interview 9, Min 29:11).

Dies könne in der Theorie sehr hilfreich sein, müsse jedoch gegen die für KI-Projekte typischerweise hohe Projektdynamik aufgewogen werden. Insgesamt erkennt der KI-Experte in Interview 9 den Methoden beider Ansätze durchaus zu, die Identifikation relevanter gesellschaftlicher wie auch ethischer Gestaltungsfaktoren systematisch zu identifizieren:

*„Ich würde es gut finden, strukturiert Nutzen und Risiken zu analysieren, diesen Raum überhaupt erst einmal aufzumachen, dass wir uns darüber unterhalten und dass es auch eine wichtige Sache ist. Natürlich unterhalten wir uns schon über Nutzen und Risiken, aber nicht so strukturiert und methodisch. Das ist ein Potenzial was man benutzen könnte“* (Interview 9, Min 37:27).

Offen bleibt für den Experten jedoch die Frage nach dem genauen Zeitpunkt der Anwendung der Methoden im Gestaltungsgeschehen. Wann in einem Gestaltungsprozess Reflexion betrieben wird und welche Bedingungen dafür erfüllt sein sollten, konkrete Reflexionsgespräche zu führen, werden ihm nicht klar genug beantwortet. Dennoch erachtet er den grundsätzlichen Vorschlag zu mehr Reflexion für sinnvoll und würde sich wünschen, dass dies künftig in den Reihen der Entscheidungsträger mehr Aufmerksamkeit erfährt. So berichtet der Experte, dass nach seiner Erfahrung die Personen, die über die Geldflüsse eines Projekts entscheiden, auch darüber entscheiden, welche Prioritäten in einem Projekt (keinen) Vorrang haben. Im Hinblick auf den Vorschlag in CTA, das Teilnehmerfeld von CTA-Workshops möglichst breit zu fächern, befindet der Experte:

*„Hier wäre eine objektivere Beurteilung [wünschenswert], mehr Leute zu befragen und mehr Antworten [zu erhalten], als wie in so einem Gespräch wo dann der lauteste und mächtigste das entscheidet“* (Interview 9, Min 38:08).

Auch in Interview 7 erläutert der KI-Experte, dass er in den Methoden durchaus ein Potenzial erkennt, sich möglichst unvoreingenommen den großen Fragen von KIGP zu widmen. Dadurch könne man bereits konkrete Werte eines Gestaltungsprojekts etablieren und kommunizieren, wenn beispielsweise durch das erhöhte Kommunikationsaufkommen, dass die Durchführung der CTA- und VSD-Methoden bedingen würde, mehr Transparenz geschaffen wird:

*„Kann ich mir schon ganz gut vorstellen, dass man High Level beginnt [...] [einzusteigen], z.B. welche Auswirkungen haben Chatbots auf die Belegschaft, [...] auf die Systemlandschaft. Das würde auf der Ebene schon ganz gut passen und ich denke da kann man sich sehr viel Transparenz erarbeiten“* (Interview 7, Min 44:10).

Darüber hinaus erkennt der KI-Experte eine Chance der Methoden im Zusammenhang mit den von ihnen angestrebten Tätigkeiten zur Folgenabschätzung. Er verwendet dazu das Beispiel der Automatisierung von Unternehmensprozessen durch den Einsatz von KI, deren Folgen mit Beteiligung verschiedener Organisationseinheiten diskutiert werden sollten:

*„Wenn es jetzt in Richtung KI und Automatisierung von bestimmten Prozessen geht, dann ist das mit Sicherheit ein Bereich wo man auch ein bisschen rechts und links schauen müsste und auch auf die Gefahren hinweisen sollte. Das sollte vor allem auch unsere Datenschutzabteilung und unseren Betriebsrat interessieren“* (Interview 7, Min 36:52).

## 6.6 Risiken

In den Risiken kommen die Bedenken der KI-Experten zum Ausdruck, die sie hinsichtlich inhaltlicher und methodischer Elemente der CTA- und VSD-Methoden geäußert haben. Risiken beschreiben Potenziale, die die Methoden in ihrer in den Interviews dargestellten Form nicht oder nur wenig thematisieren oder adressieren, die jedoch nach Einschätzung der Experten den Erfolg von KIGP gefährden können.

### 6.6.1 Prozess- und Ergebnisorientierung

In Interview 1 hinterfragt der befragte KI-Experte, inwieweit die Methoden zu dem in der praktischen KI-Gestaltung notwendigen Mindset kompatibel sind. So ist für den Experten nicht ersichtlich, wie der Einsatz der Methoden dabei unterstützen kann, die Leistungsfähigkeit von KI-Systemen zu fördern:

*„If we do this, the performance is going to decrease. [...] Once you are in this industry mindset, at least as it is nowadays, performance is everything, because it is what you benchmark against your competitors“* (Interview 1, Min 24:15).

Im Interview mit dem Experten wird deutlich, dass besonders in der Speech Recognition-Branche die Performance von KI-Systemen und das Einhalten von Deadlines zentrale Kriterien in Gestaltungsprojekten darstellen. Er äußert Bedenken, dass die CTA- und VSD-Methoden im Vergleich mit anderen Methoden zu starr seien, um seine Arbeitskollegen und ihn beim Erreichen dieser Kriterien zu unterstützen. Unterstützt wird er in seiner Einschätzung durch den KI-Experten in Interview 4. Dieser Experte bezieht sich in seinen Ausführungen primär auf die Vorschläge der CTA-Methoden und gibt zu bedenken, dass hier die Auswahl bzw. die Zusammensetzung der Workshopteilnehmer ein erfolgskritisches Kriterium darstellt:

*„Es ist halt schwierig, man muss die Leute mit dem richtigen Mindset erwischen um da was Sinnvolles rauszukriegen. Oder man schmeißt die zusammen in einen Raum und die bekommen dadurch andere Perspektiven mit und werden dadurch dann kreativer“* (Interview 4, Min 30:26).

*„Ich glaube das ist nicht so leicht, die Leute an einen Tisch zu kriegen. Inhaltlich sinnvoll [sind die Methoden] auf jeden Fall, aber vielleicht muss man auch aufpassen, dass man [sie] nur anwendet, wenn es sich lohnt. Es ist schon ein Overhead, eine große Hürde diesen Aufwand überhaupt anzufangen. Und es wäre schade, wenn ein Projekt dann scheitert“* (Interview 4, Min 33:20).

Der Experte bringt zur Sprache, dass die Anwendung der Methoden für Gestalter und Entscheidungsträger in den meisten Fällen einen Zusatzaufwand bedeutet, dessen Erbringung sich durch zielführende Beiträge und Ergebnisse zu Gestaltungszielen lohnen sollte. Ein Scheitern von Projekten aufgrund der Entscheidung, CTA- oder VSD-Methoden im Projekt anzuwenden, sei nicht vertretbar. Der Experte spricht sich deshalb dafür aus, den Einsatz der CTA- und VSD-Methoden frühzeitig und sorgfältig zu planen und möglichst zur Klärung übergreifender Fragestellungen und Zusammenhänge einzusetzen, weniger für KI-Projekte kleinen Umfangs oder Funktionen eines KI-Systems mit geringem Ausmaß. Diesem Experten fehlen demnach zusätzliche Orientierungshilfen, die Gestalter und Entscheidungsträger zu Rate ziehen und prüfen können, um sich für oder gegen einen Einsatz der Methoden entscheiden zu können. Die dadurch zum Ausdruck gebrachte, stellenweise nicht ersichtliche Beitragsfähigkeit der CTA- und VSD-Methoden wird auch von den KI-Experten der Interviews 5 und 8 bemängelt:

*„Der Ansatz mit hypothetischen Szenarien zu operieren und die dann und in einer Diskussion aufzuarbeiten – Ich weiß nicht, wie erfolgsversprechend das ist“* (Interview 5, Min 23:10).

*„Was ich mich gefragt habe bei beiden Methoden ist natürlich: Wie stellt man sicher, dass das dann auch einen Effekt hat? Also dass die [damit gewonnenen] Erkenntnisse dann auch einen Einfluss haben auf die Gestaltung. Denn es endet ja [...] bei der Erkenntnisgewinnung“* (Interview 8, Min 17:30).

Besonders aus der Aussage des Experten in Interview 8 wird deutlich, dass er den fehlenden Transfer zum aktiven Gestaltungsgeschehen als ein Risiko der CTA- und VSD-Methoden einschätzt. In seinem Beruf sei eigentlich der Schritt der danach kommt, also die Perspektive auf mögliche abbildbare Einflüsse und Veränderungen auf KI-Systeme spannend und entscheidend, dies könne er jedoch im präsentierten Zustand der Methoden nicht ausmachen. Auch dieser Experte kann dadurch nur schwer einschätzen, wie zielführend der Einsatz der Methoden in konkreten Projekten wäre. Am deutlichsten wird die mangelnde Thematisierung konkreter Resultate jedoch in Interview 9 kritisiert. Der Experte äußert hierbei besonders an den soziotechnischen Szenarien in CTA Kritik, da ihm ihr Beitrag zu den Gestaltungszielen eines KI-Projekts nicht ersichtlich wird:

*„Wenn es nicht klar ist, was das Ziel von dem Termin ist, was direkt umsetzbar ist im [Sinne von] wir kommen unserem Ziel näher, ist es immer schwer zu erklären“* (Interview 9, Min 33:10).

Hier wird erneut deutlich, dass auch diesem Experten in letzter Konsequenz die Menge an überzeugenden Argumenten für den Einsatz der Methoden in KI-Projekten fehlt, da er zu





wenig Bezüge seiner Gestaltungsanforderungen in den Methoden adressiert sieht. Auch fehlen ihm die direkt ersichtlichen Vorteile, die ein Einsatz der Methode mit sich bringt, weshalb er die Anwendung der CTA-Methode mehr als Risiko bewertet. Ein weiteres, für den KI-Experten sehr zentrales Risiko stellt der Mangel an zeitlichen Vorgaben in den Methoden dar. Der Experte beschreibt, dass eine wesentliche Anforderung für Gestalter und Entscheidungsträger in KIGP der effiziente Umgang mit Zeit sei. Dieser Aspekt finde in den Methoden beider Ansätze jedoch keinerlei Berücksichtigung, was bei dem Experten erkennbares Unbehagen weckt:

*„Ich würde mich immer fragen: Wie wäre der zeitliche Rahmen? Wie wird es genau in ein Projekt eingebettet, dass es tatsächlich nicht überhandnimmt und nicht dann schon die Motivation [...] abnimmt [...]? Dass es nicht zu sedierend wirkt“ (Interview 9, Min 38:41).*

*„Ist es ein Tag, sind es mindestens drei Tage [...]? Wie viele Stunden oder Tage es wären, für diese Methode. Für einen nicht klaren Nutzen [...]. Deshalb die Abwägung Nutzen-Kosten sozusagen, wie viel Zeit kostet uns das als ganzes Projektteam und wie viel Nutzen haben wir davon [...]. Wie rechtfertigt man das?“ (Interview 9, Min 42:10).*

Der Experte sieht in den Methoden das Risiko, besonders zu Beginn von Projekten die Euphorie und die Gestaltungskraft einzuschränken, welches seiner Erfahrung nach eine wichtige Energie- und Inspirationsquelle für alle Beteiligten in KIGP darstelle. Er beschreibt den Eindruck, dass die Methoden, wenn sie nicht zielgenau eingesetzt würden, diesem Elan zu sehr entgegenwirken würden, und dadurch das Erreichen von Gestaltungszielen mehr behindern wie fördern könnten. Auch ihm fehlen in letzter Konsequenz erkennbare und schlagende Argumente, mit deren Hilfe er den Einsatz der Methoden in KI-Projekten rechtfertigen kann.

In Interview 6 identifiziert der KI-Experte ein potenzielles Risiko der Methoden darin, auf methodischer Ebene zu unflexibel zu sein. Agilität sei in KIGP zwingend erforderlich, was notwendig mache, im Verlauf des Prozesses auf methodischer Ebene Veränderungen durchzuführen:

*„Das sollte so sein wie Softwareentwicklung – agil [...]. Man darf sich nicht jetzt an dieser Methode festsetzen und sagen [...] ich will ein Ergebnis haben, obwohl man merkt, [...] da läuft was falsch. [Dann] wird [es] politisch, ich habe eine Sache angefangen und kann nur nicht sagen, dass es falsch ist. Muss es bis zum Ende [durch]ziehen. Das ist gefährlich, weil dann kommen Ergebnisse raus, die man am Ende gar nicht braucht“ (Interview 6, Min 27:25).*

Er warnt explizit davor, Gestaltungsmethoden als politisches Instrument zu betrachten. In Gestaltungsprozessen sei es regelmäßig erforderlich, getroffene Entscheidungen einer aufmerksamen und ehrlichen Überprüfung auf Zielführung zu unterziehen, und bei Schlussfolgerungen, die gegenteilige Ergebnisse liefern, Korrekturmaßnahmen einzuleiten. Ein weiteres übergreifendes Risiko spricht der KI-Experte in Interview 2 an. Er bringt

zu Sprache, dass nach seiner Erfahrung viele Kunden und Auftraggeber auf die Anschaffung und den Einsatz von KI-Systemen dieselben Kriterien anlegen, wie auf nicht-digitale Produkte oder digitale Systeme, die gänzlich andere Zwecke erfüllen sollen:

*„Das ist [...] einer der wichtigsten Punkte, [...] dass alle in diesen Projektteams oder die in diesen Gruppen beteiligt sind, dass die [...] ein gleiches Verständnis dafür haben, was genau die KI jetzt eigentlich ist und was sie auch damit machen wollen.“ (Interview 2, Min. 20:00)*

*„Ich glaube das ist einigen Leuten nicht bewusst [...]. Es ist einfach immer noch ein extremes Risiko, es ist etwas Anderes [wenn man] ein Stück Stahl [einkauft] als [wenn] man eine KI [einkauft] und möchte die anwenden. Die sehen das nicht kommen, dass da ein Risiko drin ist.“ (Interview 2, Min 22:58)*

Er kritisiert an den Methoden von CTA und VSD, dass sie nicht in der Lage seien, die sehr spezifischen Merkmale und Funktionen von KI-Systemen kommunizieren zu können und die damit verbundenen Chancen und Risiken mit Kunden auf Augenhöhe zu diskutieren. Ihm fehlt an den Methoden die Möglichkeit, spezifisches Wissen zu KI und zur KI-Gestaltung im Kontakt mit Stakeholdern in den Vordergrund zu rücken und so ein möglichst gemeinsames Verständnis von KI-Systemen und deren Zwecken zu entwickeln. Ein solches Merkmal ordnet er seiner Auffassung von einer menschenzentrierten Gestaltung zu, weshalb die CTA- und VSD-Methoden unter dem Gesichtspunkt der Schaffung von Transparenz und Kommunikationsfähigkeit gegenüber Stakeholdern unzureichend seien. Darin klingt auch an, dass er an den Methoden beider Ansätze kritisiert, dass sie potenzielle Anwender nicht dazu befähigen würden, Anwendungsfälle für KI-Systeme zu identifizieren.

## 6.6.2 Reflektierter Umgang mit Verantwortung und Befangenheit

Unter dem Gesichtspunkt, dass beide Ansätze anstreben, durch ihre Methoden Gestalter und Entscheidungsträger zur (stärkeren) Reflexion zu befähigen, werfen zwei der KI-Experten berechnete Fragen auf. In Interview 6 bezieht sich der KI-Experte auf die Personen in der Rolle der CTA agents, die durch ihre fachlichen Expertisen Vorlieben und Aversionen aufweisen würden, die sich auf ihre Arbeit und entsprechend auch auf KI-Gestaltungsprozesse übertragen können:

*„Die Frage ist natürlich: Sind CTA agents geprimet mit gewissen Themen oder Vorlieben oder Aversionen gegenüber bestimmten Themen? Das ist natürlich immer schlecht. Das kann sein, dass sie bestimmte Entwicklungen, aus welchen Interessen auch immer, lenken können“* (Interview 6, Min 30:15)

Er verweist hier auf den Priming-Effekt, der das Phänomen einer selektiven Wahrnehmung beschreibt, die interessenskonforme Themen aufgreift und anderslautende Aspekte teilweise bewusst, teilweise unbewusst ausblendet. Aus dieser Aussage wird auch deutlich, dass der Experte den in CTA und VSD zum Ausdruck gebrachten Anspruch der Reflexionsförderung auf die Methoden selbst anwendet, und nach Maßnahmen fragt, die neben den Gestaltern und Entscheidungsträgern auch die Prozessbegleiter dazu auffordert, sich mit ihren möglicherweise unbewusst vorhandenen Interessen und Agenden auseinanderzusetzen:

*„Wenn jemand versucht [Gestaltungsprozesse] zu moderieren, [... eine] kontrollierte Umgebung [zu schaffen ...], kann [er] die Leute manipulieren. Und das ist eine Gefahr die man hier hätte. Und da muss man natürlich dagegenwirken“* (Interview 6, Min 31:10).

Für diesen Experten bedeutet Moderation immer auch das Ausüben einer Form von Kontrolle, weshalb er in CTA in der Rolle der CTA agents ein Risiko zur Manipulation identifiziert, dem gegenüber keine Schutzvorkehrungen getroffen sind. Dementsprechend vermisst er in den CTA-Methoden, dass sie auch die CTA-agents explizit zu einem reflektierten Umgang mit der Verantwortung aufgefordert werden, die das Ausüben ihrer Rolle mit sich bringt. Auch für den KI-Experten in Interview 5 ist die Art und Qualität der Moderation in den CTA-Workshops ein wesentliches Kriterium, der für den Erfolg oder Misserfolg der Methode ausschlaggebend ist:

*„Wie läuft die Moderation ab?“* (Interview 5, Min 33:22).

Das darin angesprochene Risiko eines unreflektierten Umgangs mit möglicherweise vorhandenen Befangenheiten kommt auch in Interview 8 zur Sprache. Der Experte beschreibt dabei seinen Erfahrungswert, dass Gestalter und Entscheidungsträger sich in der Regel mit einem Projekt identifizieren und deshalb ein hohes Eigeninteresse haben, ein Gestaltungsprojekt zum Erfolg zu führen. Die VSD-Methode der Stakeholderanalyse würde jedoch Gestalter und Entscheidungsträger nicht dazu auffordern, die im Projekt vertretenen

Werte und vorhandenen Interessen auf dieselbe Weise zu analysieren, wie die der Stakeholder:

*„Das Projektteam hat natürlich ein Interesse daran, sich immer mit dem Projekt [zu] identifizieren und möchte das zum Erfolg führen. Es hat nicht so das Incentive, Risiken, die sie selbst vielleicht nicht so betreffen [...] mit derselben Sorgfältigkeit auszuarbeiten oder genauso stark zu gewichten, wie das vielleicht bei einem unabhängigen Agent in der anderen Methode der Fall wäre“ (Interview 8, Min 29:23).*

*„Das man halt befangen ist, wenn man mit dem Projekt verbunden ist. Wenn man keine objektive [Perspektive] von außen, noch eine Begleitung hat.“ (Interview 8, Min 37:00).*

Auch wenn dieser Experte dem Experten aus Interview 6 in Bezug auf CTA widerspricht, bringt auch er den Aspekt der Befangenheit zu Sprache, dem ihm in der VSD-Methode zu wenig Aufmerksamkeit gewidmet wird.

### 6.6.3 Partizipation von Stakeholdern

In Ergänzung zu den Stärken und Schwächen identifizieren KI-Experten in zwei Interviews zusätzliche Risiken von Partizipationsformaten, wie sie in den CTA- und VSD-Methoden Anwendung finden. So zweifelt der Experte in Interview 8 daran, dass die CTA- und VSD-Methoden geeignete Werkzeuge darstellen, um Chancen und Risiken von KI zu ermitteln:

*„Ich glaube dann ist es eine Challenge, neutral darzustellen, also die Chancen und Risiken von KI. Also die so darzustellen, dass sie richtig eingeordnet werden von diesen Gruppen, die in den Workshops teilnehmen“ (Interview 8, Min 33:38).*

Er begründet seine Bedenken mit dem Hinweis auf die zeitgenössische Medienberichterstattung zum Thema KI, die einerseits „Wunder“ verspreche, andererseits „Horrorszenarien“ aufzeige. Deshalb sei es in der Interaktion mit Stakeholdern wichtig, Informationen über die Möglichkeiten und Grenzen von KI verlässlich und verständlich transportieren zu können. In Bezug auf die VSD-Methoden äußert sich zudem der KI-Experte aus Interview 9 skeptisch, da er ein Risiko darin sieht, (fachfremde) Stakeholder Chancen und Risiken von KI identifizieren zu lassen:

*„Dass man das auf der Ebene von einzelnen Vertriebsmitarbeitern direkt identifizieren kann, das fanden die Stakeholder richtig klasse, das hätten die richtig gut gefunden [...]. Ich meine nur, dass ich nicht weiß ob das die richtigen Leute sind, die das beurteilen, die Stakeholder, weil die eigene Interessen haben“ (Interview 9, Min 25:38).*

## 6.7 Diskussion

Die Einschätzungen der KI-Experten zu den CTA- und VSD-Methoden lassen auch für die Debatte zur Entwicklung von Gestaltungsmethoden in HCAI relevante Schlussfolgerungen zu. Die Einschätzungen zu den Stärken der CTA- und VSD-Methoden zeigen, unter welchen Gesichtspunkten sich die Methoden auch für eine Erwägung in HCAI eignen (Forschungsfrage 4). Hierbei sind für die Experten besonders die Eigenschaften bedeutsam, durch die Methoden Gestaltungs- und Managementaufgaben miteinander zu verbinden, die gestalterische Kreativität zu fördern und in Kooperation mit Stakeholdern Lösungen zu ethischen und gesellschaftlichen Fragestellungen zu erarbeiten. Die identifizierten Schwächen und Risiken zeigen auf, was die Methoden für den HCAI-Diskurs nicht bereitstellen können (Forschungsfrage 5): Sie machen keine konkreten Angaben zu für die Durchführung einzelner Methodenschritte benötigten Zeitkontingenten oder zu der für die Gestaltungspraxis besonders bedeutsame Rückführung erarbeiteter Erkenntnisse in das Gestaltungsgeschehen. Auch identifizieren die KI-Experten einige in den Methoden vorhandene Biases, wenn beispielsweise in VSD die für Systeme verantwortlichen Gestalter und Entscheidungsträger selbst die Chancen und Risiken der von ihnen gestalteten Systeme identifizieren sollen. Die KI-Experten identifizieren darüber hinaus verschiedene Chancen, die Hinweise beinhalten, wie die Praxisauglichkeit der CTA- und VSD-Methoden in KIGP künftig weiter gefördert werden kann (Forschungsfrage 6). Hierzu zählen die Möglichkeiten, mithilfe der Methoden Technikfolgen systematisch abzuschätzen, Wissen über KI zu vermitteln oder auch um bei Gestaltern und Entscheidungsträgern Unsicherheiten auszuräumen. In Summe entsteht so ein differenzierter Eindruck über die Eignung der CTA- und VSD-Methoden für die Gestaltung von KI-Systemen, die die KI-Experten anhand verschiedener Kriterien beurteilen. Die Mehrzahl der von den KI-Experten identifizierten Stärken, Schwächen, Chancen und Risiken sprechen für sich, hinsichtlich der Fragestellung der Dissertation erfordern jedoch verschiedene der aufgezeigten Aspekte eine detailliertere Auseinandersetzung. Diese Aspekte diskutiere ich nachfolgend und beschreibe, inwieweit sie für den MZKG-Diskurs in HCAI Bedeutung erfahren. Zu Beginn reflektiere ich zudem kurz über die Funktion sowie über die Inhalte der Storyboards in den Interviews, da sie eine Limitierung der Interviewstudie darstellen.

### 6.7.1 Zur Funktion und den Inhalten der Storyboards

Den Storyboards kommt in den Interviews eine wesentliche Funktion zu, da die KI-Experten sich zu großen Teilen auf die darin bildlich und schriftlich dargelegten Informationen beziehen, um ihre Einschätzungen abzuleiten und ihre Beurteilungen zu formulieren. Bei der Entwicklung der Storyboards habe ich mich stark an der Literatur zu beiden Ansätzen orientiert, allerdings werden auch dort relevante Details zu den Methoden ausgespart. Dabei handelt es sich im Besonderen um die von den KI-Experten vermissten Informationen zu Zeitumfängen für bestimmte Arbeitsschritte und näheren Informationen über konkrete Formate von Workshops oder Fragebogenlayouts. Eine weitere Einschränkung der Storyboards liegt in der Auswahl von VSD-Methoden, die ich für den Rahmen der

Interviewstudie treffen musste. Zwar konnte ich über alle Interviews hinweg die tripartite methodology abbilden, jedoch war es mir nicht möglich, im Rahmen der Expertenevaluation den vollständigen Korpus von 17 Methoden zu evaluieren, der in der VSD-Literatur ausgewiesen wird. Darüber hinaus habe ich in der Entwicklung der Storyboards eine didaktische Reduktion der CTA- und VSD-Methoden vorgenommen, um während der Interviews bei den KI-Experten eine möglichst geringe kognitive Belastung zu erzielen und ihre Aufmerksamkeit auf die wesentlichen Aspekte der Methoden auszurichten. Durch die didaktische Reduktion besteht jedoch die Möglichkeit, dass in der CTA- und VSD-Literatur beschriebene Details zu den Methoden verloren gehen, und deshalb in den Interviews nicht thematisiert werden. So stelle ich in den Storyboards die Bedeutsamkeit einzelner Elemente der Methoden weniger prominent dar, wie sie in der Literatur zu CTA und VSD dargestellt werden. Ein konkretes Beispiel für meine im Rahmen der didaktischen Reduktion durchgeführten Anpassungen ist der Begriff des „Projektteams“, den ich in der Beschreibung der VSD-Methoden häufig verwende um darzustellen, welche Gruppen und Individuen Tätigkeiten durchführen. Dieser Begriff wird in dieser expliziten Form nicht in der VSD-Literatur verwendet. Insofern sind bei der Interpretation der Erkenntnisse der Expertenevaluation diese Limitierungen zu berücksichtigen.

### 6.7.2 Transfer von Reflexionserkenntnissen

Die Förderung von Reflexion bei Gestaltern, Entscheidungsträgern und Stakeholdern ist eines der bedeutsamsten Ziele der Methoden von CTA und VSD. Die CTA- und VSD-Methoden widmen sich diesem Ziel, indem sie verschiedene Instrumente der qualitativen und quantitativen Forschung aufgreifen und adaptieren, oder wie im Falle von VSD, auch gänzlich neue Methoden wie beispielsweise die Stakeholder tokens entwickeln. Das Schaffen von Reflexionsräumen ist auch für die Ziele der MZKG-Debatte in HCAI insofern zielführend, da sich das Feld durch KI, im Vergleich mit Softwaresystemen, mit einer deutlich komplexeren Technologie konfrontiert sieht, und deren menschenzentrierte Gestaltung deutlich komplexere Fragestellungen zu ethischen und gesellschaftlichen Sachverhalten aufwirft. Wie aus den Interviews hervorgeht, können es die Methoden von CTA und VSD ermöglichen, das Bewusstsein aller am Gestaltungsprozess beteiligten Gruppen und Individuen zu schärfen, und es auf die bedeutsamen Aspekte solcher sozialen und ethischen Fragestellungen von KIGP auszurichten. Die Methoden sind in der Lage, gedankliche Räume zur Identifikation, Analyse und Diskussion von Gestaltungsfaktoren zu schaffen, wenn sie beispielsweise vorsehen, Projektteams mit der Identifikation zentraler Werte eines Gestaltungsprojekts zu beauftragen, oder Technologiefeldakteure in CTA-Workshops zur Diskussion soziotechnischer Szenarien aufzufordern. Durch die Formulierung präziser Aufgaben-, Frage- oder Problemstellungen, die in solchen Formaten diskutiert werden, kann auch in HCAI sichergestellt werden, dass relevante Antworten auf bedeutsame Gestaltungsfragen der MZKG ermittelt werden. Die Betonung muss hierbei jedoch auf der Formulierung *präziser* Fragen liegen, denn wie ebenfalls aus den Interviews hervorgeht, muss in der Gestaltungspraxis möglichst vor der Anwendung von Gestaltungsmethoden ihr Beitrag zu den Gestaltungszielen erkennbar sein. Die Entwicklung von beispielsweise soziotechnischen Szenarien oder Prototypen müssen Zwecken dienen,

die eine hohe Relevanz in einem KI-Gestaltungsprojekt besitzen. Einer der stärksten Kritikpunkte der KI-Experten bezieht sich auf die mangelhaften Vorkehrungen und Hinweise in den Methoden um sicherzustellen, dass ein Transfer der mit den Methoden erarbeiteten Erkenntnisse in das tatsächliche Gestaltungsgeschehen erfolgt. Dieser Kritik kann auch im Feld HCAI begegnet werden, indem in künftigen Ausbaustufen der CTA- und VSD-Methoden interessierten Anwendern Detailangaben wie beispielsweise konkrete Maßnahmenkataloge zur Überführung von Workshop- oder Fragebogenergebnissen an die Hand gegeben werden. Vorstellbar ist auch, hierfür Strategien und Erfahrungswerte aus Feldern wie dem Requirements Engineering oder dem User Research aufzugreifen. Eine weitere Möglichkeit besteht darin, interessierten CTA- und VSD-Anwendern Beispiele an die Hand zu geben, die illustrieren, wie Methoden und daraus resultierende Ergebnisse in der Vergangenheit in (KI-)Gestaltungsprozesse integriert wurden. Hierbei gilt es jedoch, wie es in den Aussagen mehrerer KI-Experten zum Ausdruck kommt, darauf zu achten, keine ausschließlich wissenschaftlichen Projektbeispiele zu verwenden („Success Stories“; siehe Kap. 6.4.1). Durch solche Hilfsmittel können interessierte Anwender in der Phase, in der sie die Tauglichkeit einer Methode für ihr konkretes Gestaltungsprojekt prüfen, besser abschätzen, welche Art von Ergebnissen sie durch den Einsatz einer Methode erwarten können. Darin liegt ein Mehrwert, um die für die MZKG sehr relevanten Gedanken und Ideen in CTA und VSD auch in der Gestaltungspraxis stärker zu verankern. In der VSD-Literatur machen Friedman und Hendry (2019) einen ersten Aufschlag in diese Richtung, da sie für jede VSD-Methode den Entwicklungskontext erläutern und auf beispielhafte Anwendungsfälle verweisen, die in der Vergangenheit in der Literatur dokumentiert wurden. Doch auch hier beziehen sich die wenigen Beispiele auf überwiegend wissenschaftliche Projektkontexte, die bei genauerer Betrachtung sehr sparsam mit verständlichen Details zur konkreten Methodenanwendung umgehen.

### 6.7.3 Steigerung von Ausbildungsaktivitäten

Die Herausforderung der Entwicklung von Gestaltungsmethoden in HCAI repräsentiert die Aufgabe, die theoretisch erarbeiteten Erkenntnisse zur MZKG in die Gestaltungspraxis zu transferieren und dort zu evaluieren. Für die HCAI-Community gilt es, eine Brücke zwischen Theorie und Praxis zu schaffen, durch die es möglich wird, aus den erarbeiteten theoretischen Erkenntnissen in der Praxis Nutzen zu ziehen. Auf die Debatten zur MZKG müssen Taten folgen, um Gestaltungspraktiker bei der Realisierung menschenzentrierter KI-Systeme zu unterstützen. Wie aus einem der Experteninterviews hervorgeht, besteht eine Möglichkeit zur Unterstützung und Förderung dieses Erkenntnistransfers darin, nachfolgende Generationen von KI-Gestaltern vermehrt in den Denkweisen in HCAI auszubilden und sie mit den Zielsetzungen der MZKG vertraut zu machen.

Diese Möglichkeit kommt in Interview 1 zur Sprache, wo der befragte KI-Experte die Konzepte von intermediate gates und virtuous cycles anspricht, die die CTA- und VSD-Methoden aus seiner Sicht erfüllen können:

*„People in industry are maybe more aiming at monetizing goals. They are never going to think about [societal or ethical issues]. I think it's important that this starts to be on surface of the society, so the new generations have it in the mindset and not in the law. Because the law will [only] be for the minimum“* (Interview 1, Min 34:00).

Er sieht einen Bedarf dafür, dass auf gesellschaftlicher Ebene ein Wandel stattfindet, sodass künftige Gestalter und Entscheidungsträger sich aus eigenem Interesse mehr mit den sozialen und ethischen Implikationen von KI-Systemen und ihrer Gestaltung auseinandersetzen. Diese Funktion können, nach Ansicht des KI-Experten, die CTA- und VSD-Methoden auch in HCAI erfüllen, wenn sie an den Universitäten in entsprechenden Studiengängen aufgegriffen und diskutiert würden. Konkret ist beispielsweise vorstellbar, dass künftige HCAI-Curricula Praxisprojekte vorsehen, in denen unter Anwendung der CTA- und VSD-Methoden konkrete Fragestellungen der MZKG-Gestaltungspraxis von künftigen KI-Gestaltern exploriert und bearbeitet werden. Auch stellen die Gedankenmodelle der intermediate gates und der virtuous cycles wertvolle Ressourcen dar, die in der Entwicklung und besonders auch Vermittlung von HCAI-Methoden aufgegriffen werden können. In der gezielten Aufnahme der CTA- und VSD-Methoden in moderne HCAI-Ausbildungsprogramme (siehe Kap. 3.2.1) besteht somit für die HCAI-Community eine weitere Möglichkeit, die Beitragsfähigkeit der Methoden für die Ziele in MZKG zu evaluieren. Ein begrüßenswerter Nebeneffekt wäre, dass nachrückende Generationen von KI-Gestaltern mit den Denkweisen und Methoden sozial- und wertorientierter Gestaltung vertraut werden.

#### 6.7.4 Aufgreifen bewährter Mixed-Method-Ansätze

Ein weiterer Aspekt der in den Interviews zur Sprache kommt, ist die Kombinationsfähigkeit der CTA- und VSD-Methoden. Die Tatsache, dass in beiden Ansätzen auf etablierte und validierte wissenschaftliche Methoden zurückgegriffen wird, wird von den KI-Experten begrüßt und auch für die KI-Gestaltung als geeignet eingeschätzt. Einige der Experten haben mit vergleichbaren Methoden in konkreten Projekten bereits positive Erfahrungen gemacht und zielführende Ergebnisse erzielt. Die Kombination der Methoden beider Ansätze ermöglicht es, qualitative Methoden (z.B. Workshops, Interviews) und quantitative Methoden (z.B. Fragebögen, Umfragen) zu kombinieren, wie sie besonders auch in der tripartite methodology in VSD Anwendung findet. Solche Mixed-Method-Ansätze können dabei helfen, die verschiedenen konzeptionellen wie auch methodischen Fragestellungen zu beantworten, die im MZKG-Diskurs in HCAI im Vordergrund stehen und deren Charakteristika es erfordern, mit qualitativem wie auch quantitativem Datenmaterial zu arbeiten. Durch die Kombination können die Stärken beider Verfahrenstypen in HCAI dazu genutzt werden, um beispielsweise die Chancen und Risiken von KI-Systemen in einer Domäne zu identifizieren (qualitative Methoden), um sie später im konkreten Nut-



zungskontext zu validieren (quantitative Methoden). Eine zielführende Begegnung der Herausforderungen der MZKG erfordert die Kombination beider Forschungs- und Gestaltungstraditionen. Über die Potenziale und Vorzüge solcher Kombinationen sind in der HCI-Community bereits Erfahrungswerte vorhanden, die entsprechend in der HCAI-Community aufgegriffen und weiterentwickelt werden können.

### 6.7.5 Gestaltung und Projektmanagement

Die Methoden von CTA und VSD wurden in den 1990er- und 2000er-Jahren vor dem Hintergrund entwickelt, Gestalter von technischen Systemen bei ihrer Arbeit zu unterstützen und sie zur Berücksichtigung sozialer und ethischer Fragestellungen zu befähigen. Entsprechend wird in den CTA- und VSD-Methoden mit traditionellen Gestalter- und Nutzerrollen gearbeitet und viele der Rollen ausgeblendet, die sich mit organisatorischen Aufgaben und Fragestellungen von Gestaltungsprozessen befassen. Wie die Expertenevaluation der Methoden gezeigt hat, werden in der KI-Gestaltungspraxis jedoch Gestaltungstätigkeiten häufig im Verbund mit Tätigkeiten des Projektmanagements ausgeführt, weshalb die Methoden beider Ansätze unter diesem Gesichtspunkt Lücken und Defizite aufweisen – Aufgaben und Tätigkeiten in Bereichen wie der Projektplanung, Koordination, Ressourcenallokation, Ressourcenzuteilung oder strategischen Entscheidungsfindung werden in CTA und VSD wenig bis gar nicht berücksichtigt. Auch wenn die KI-Experten vereinzelt Aussagen treffen, dass die Methoden auch zu Aufgaben des Projektmanagements komplementär sind, sind die CTA- und VSD-Methoden keine Projektmanagementtools, wie sie in der MZKG benötigt werden. Das sollten sie nie sein. Die HCAI-Community sieht sich im MZKG-Diskurs jedoch mit der Herausforderung konfrontiert, Methoden und Frameworks zu entwickeln, die es Gestaltern und Entscheidungsträgern ermöglichen, Aufgaben aus beiden Tätigkeitsfeldern durchzuführen – Methoden zur MZKG müssen Gestaltungs- *und* Projektmanagementtools sein, oder zumindest Schnittstellen zu Tätigkeitsfeldern des Projektmanagements aufweisen, wenn die Methoden dem Anspruch an Praxistauglichkeit gerecht werden sollen. Unter diesem Gesichtspunkt können die Methoden von VSD und CTA jedoch keine Beiträge zur Methodenentwicklung in HCAI leisten, da in ihnen Aspekten wie der Umgang mit knappen Zeitbudgets oder den Herausforderungen des Ressourcenmanagements kaum Beachtung geschenkt wird.

### 6.7.6 Klärung von Verantwortungsbereichen

Damit verbunden ist ein weiterer tiefgreifender Aspekt, der gegenwärtig die fachbereichsübergreifende KI-Debatte beschäftigt. Es handelt sich dabei um ein vorhandenes Vakuum zur Frage der Verantwortung und Verantwortbarkeit: Welche Rollen in KIGP können und sollen sinnvoller und angemessener Weise dafür verantwortlich gemacht werden, dass in KIGP ethische und soziale Aspekte adressiert werden? Diese zentrale Frage ist auch in HCAI Gegenstand im Diskurs zur MZKG, mit deren Auseinandersetzung jedoch sehr weitreichende und sehr komplexe Zusammenhänge verknüpft sind. Auch die HCAI-Community konnte deshalb noch keine zufriedenstellenden Antworten

erarbeiten. Bei der Gestaltung technisch simplerer Softwaresysteme fiel die Klärung dieser Verantwortungsfrage leichter – so besagt der etablierte Konsens (beispielsweise auch in der MZG), dass die Gestalter eines Systems dafür Sorge zu tragen haben, dass Aspekte der Mensch- und Werteorientierung in Gestaltungsprozessen angemessen berücksichtigt und adressiert werden. Allerdings war der damit assoziierte Verantwortungsbereich deutlich überschaubarer wie jener der KI-Gestaltung, da die Einsatzkontexte von Softwaresystemen häufig recht klar definiert werden können, was auch die Antizipation der mit ihrem Einsatz verbundenen Folgen erleichtert. In der KI-Gestaltung hingegen stellt sich die Sachlage weitaus komplexer dar, was bereits aufgrund der Tatsache, dass KI-Systeme weit vielfältigere Einsatzmöglichkeiten eröffnen, ersichtlich wird. Ähnlich wie beim Diskurs zu Softwaresystemen wendet sich der gesellschaftliche und wissenschaftliche Diskurs zu Fragen der Verantwortung und Verantwortbarkeit ebenfalls recht schnell der Gestaltung der KI-Systeme zu, und bringt dieselben Forderungen an: In KIGP sollen ethische Werte eingehalten werden und zu technischen Lösungen führen, die einer Prüfung nach in der Gesellschaft verankerten ethischen Gesichtspunkten standhalten müssen und ethische Grundsätze nicht verletzen sollen. Dies führt jedoch, wie es Jobin et al. treffend formulieren, in Feldern wie der Ethical AI und der HCAI unweigerlich zu der Frage:

*„what constitutes ‘ethical AI’ and which ethical requirements, technical standards and best practices are needed for its realization“ (Jobin et al., 2019, S. 389).*

Die Flut an Ethik-Richtlinien, die in den vergangenen Jahren mit Blick auf die KI-Gestaltung entwickelt wurden, geben Aufschluss zur inhaltlichen Dimension ethischer KI, helfen aber in der Gestaltungspraxis wenig, um die aus diesen Forderungen resultierenden, tatsächlichen Probleme zu lösen. Und wie in Interview 1 zum Ausdruck kam (siehe Kap. 6.7.3), kann die Sicherstellung ethischer Standards durch rechtliche Vorgaben nur ein Minimum der ethischen Herausforderungen von KIGP adressieren. VSD hingegen entspringt einer Forschungstradition, die sich sehr pragmatisch mit den ethischen Fragen von Gestaltungsprozessen auseinandersetzt. Die KI-Experten bringen in den Interviews unter diesem Gesichtspunkt mehrfach ihre Anerkennung und Dankbarkeit zum Ausdruck, da sie erkennen, dass in den VSD-Methoden Versuche unternommen werden, besonders den Anforderungen von Gestaltern und Entscheidungsträgern Rechnung zu tragen. Auch wenn die VSD-Community gegenwärtig noch erst am Anfang steht, sich im KI-Diskurs zu positionieren, kann sie nach meiner Einschätzung künftig dabei helfen zu klären, welche Rollen für die Einhaltung der Mensch- und Werteorientierung in der MZKG angemessener Weise verantwortlich gemacht werden können. Für die HCAI-Community kann es sich deshalb unter diesem Gesichtspunkt lohnen, die VSD-Methoden aufzugreifen und unter den spezifischen Anforderungen der MZKG weiterzuentwickeln.

### 6.7.7 Flexibilität und Agilität

Bezüglich der Flexibilität der CTA- und VSD-Methoden zeigen die KI-Experten in den Interviews geteilte Ansichten. Einerseits begrüßen manche Experten an den Methoden, dass interessierte Anwender mit ihnen steuern können, wie viel Zeit und Ressourcen Gestalter und Entscheidungsträger für die Identifikation und Abwägung von Gestaltungsfaktoren aufbringen müssen. Andere Experten kritisieren hingegen die Starrheit der Methoden. Die mehr kritischen Stimmen sprechen damit das Merkmal der Agilität an, die in der IT-Gestaltungspraxis von zentraler Bedeutung ist und deshalb auch für die MZKG Relevanz erfährt. Gestalten und Entscheidungsträgern muss es möglich sein, im Gestaltungsgeschehen auf unvorhergesehene Entwicklungen oder Herausforderungen reagieren zu können, weshalb die dafür erforderliche Flexibilität häufig bereits in die Projektplanung miteinfließt. Auf inhaltlicher Ebene erfüllen die CTA- und VSD-Methoden die Kriterien der Flexibilität und Agilität, da sie Gestalten und Entscheidungsträgern ermöglichen, die individuellen Anforderungen von Gestaltungsprojekten Rechnung tragen können. Die von manchen KI-Experten geäußerte Kritik einer fehlenden Agilität der Methoden führe ich auf meine Darstellung der Methoden in den Storyboards zurück, die ich aus Gründen der didaktischen Reduktion in Form linearer (nicht zyklischer) Handlungsschritte dargestellt habe. Dies weckt bei den diesen Aspekt kritisierenden KI-Experten erkennbares Unbehagen, da sie in ihren Professionen das Denken und Handeln in rasch aufeinanderfolgenden, kurzen Zeitzyklen verinnerlicht haben. In der CTA- und VSD-Literatur wird nicht explizit ausgeschlossen, dass die Methoden anstelle einer linearen Durchführung auch iterativ durchgeführt werden können. Es ist durchaus vorstellbar, dass mehrere Iterationsschleifen von beispielsweise CTA-Workshops oder Befragungen zu Value dams and flows durchgeführt werden können, womit die Methoden der beiden Ansätze auch dem MZG-Gestaltungsprinzip der Iteration (siehe Kap. 2.3.3) Rechnung tragen können. Dennoch ist die von den KI-Experten geäußerte Kritik mangelnder Agilität ernst zu nehmen, da sie zeigt, welch hohen Stellenwert dieses Kriterium in der KI-Gestaltungspraxis für die Evaluation von Gestaltungsmethoden einnimmt.

### 6.7.8 Verankerung der Folgenabschätzung in der Praxis

Die aktive Auseinandersetzung mit intendierten und nicht-intendierten Effekten von KI-Systemen ist ein zentrales Themenfeld der MZKG-Debatte in HCAI. Aspekte wie der Schutz von Privatsphäre, geistigem Eigentum oder die angemessene Förderung von menschlichen Kooperationsprozessen werden aufgrund der Möglichkeiten der KI-Technologie stellenweise unter völlig neuen Gesichtspunkten diskutiert und erfordern von Gestaltern und Entscheidungsträgern ausgeprägte Mindsets, um die vermeintlich wünschenswerten Effekte gegen potenziell nachteilige Effekte abzuwägen. Wie die Experteninterviews verschiedentlich gezeigt haben, sind sich die KI-Experten in der Gestaltungspraxis ihrer dahingehenden Verantwortung bewusst und befinden darüber hinaus die CTA- und VSD-Methoden für tauglich, die Abschätzung von Technikfolgen in konkreten Gestaltungsprozessen einzubetten und durchzuführen. Ein KI-Experte assoziiert die Methoden dabei beispielsweise mit den (auch in TA-Arbeiten oftmals ausgewiesenen) Funktionen eines „Frühwarnsystems“, da die Methoden bereits in der Projektplanung und

zu Projektbeginn das Anstellen von Gedanken zu möglichen Chancen und Risiken vorsehen. Allerdings geben die KI-Experten auch zu bedenken, dass die dafür notwendigen ganzheitlichen Perspektiven auf das Gestaltungshandeln aufgrund konkreter Herausforderungen in KIGP erfahrungsgemäß verloren gehen (siehe Kap. 6.4.2). Aus diesem Grund kann es ein Mehrwert für die künftigen Prozessmodelle der MZKG darstellen, in ihren zyklisch angelegten Verfahren Gestaltern und Entscheidungsträgern die Möglichkeit zu eröffnen, mögliche Folgen der von ihnen gestalteten Systeme *kontinuierlich* zu beachten und zu evaluieren.

## 6.8 Zusammenfassung

Die Expertenevaluation der Methoden von CTA und VSD hat gezeigt, welche Anforderungen in der KI-Gestaltungspraxis an Gestaltungsmethoden gestellt werden. Die KI-Experten stimmen mit der in HCAI vertretenen Position überein, dass die verstärkte Auseinandersetzung mit ethischen und sozialen Implikationen über das Jahr 2023 hinaus auch in der praktischen KI-Gestaltung ein zentrales Aufgabenfeld darstellt. Jedoch weisen die Experten auch darauf hin, dass es für ein Erreichen von Gestaltungszielen unter realen Bedingungen in KIGP erforderlich ist, Informationen zu Aspekten wie benötigtem Ressourcenaufwand oder vorstellbaren Ergebnissen bereitzustellen. Die Experteninterviews haben ferner gezeigt, wie sich diese Priorisierung im Gestaltungsalltag für KI-Experten äußert, welche Denkmuster sie dadurch etabliert haben, und welche Kriterien und Argumentationsstrategien sie anwenden, um die CTA- und VSD-Methoden zu beurteilen. Im direkten Vergleich zwischen CTA und VSD favorisiert die Mehrheit der KI-Experten die Methoden von VSD, da die VSD-Methoden inhaltlich und formell viele Überschneidungen zu den für KI-Experten üblichen Anforderungen aufweisen. Den KI-Experten entgeht nicht, dass in VSD durch die Verwendung des Projektbegriffs und -gedankens eine im Vergleich mit CTA niedrigere Abstraktionsebene zum Tragen kommt, wenn über Prozesse der TG nachgedacht und gesprochen wird. Dadurch dass sich in VSD auf Kooperationszusammenhänge in Projektform bezogen wird, können die Experten die vorgeschlagenen Methoden einfacher einordnen und gedanklich auf vorstellbare Situationen ihres Berufslebens übertragen. Dies fiel ihnen bei den Methoden von CTA tendenziell schwerer, da ihnen die CTA-Methoden (noch) weniger Orientierungsgrößen und konkrete Rahmenvorgaben an die Hand geben. So hätte sich die Mehrzahl der Experten mehr Informationen zu Aspekten gewünscht, wie der Anzahl von Organisationen oder Akteuren, die in einer Technologiefeldanalyse mindestens analysiert werden sollten, oder ob bereits mit der Durchführung eines einzigen CTA-Workshops zielführende Erkenntnisse und Ergebnisse erzielt werden können. In Summe wird erkennbar, dass die CTA- und VSD-Methoden nicht dediziert für die KI-Gestaltung entwickelt wurden, dass sie jedoch durch verschiedene Anpassungen und Konkretisierungen anschlussfähig zu den methodischen Herausforderungen der MZKG-Debatte sind.

## 7 Folgerungen

Die Konzeption und Umsetzung von Gestaltungsprozessen für menschenzentrierte KI-Systeme erfordert die Integration und Berücksichtigung eines umfassenden Kriterienkatalogs, der technische, organisatorische, soziale und auch ethische Gestaltungsfaktoren beinhaltet. Die HCAI-Community hat sich zum Ziel gesetzt, Lösungen für diese Herausforderung zu erarbeiten und Gestaltungsframeworks zu entwickeln, die Gestaltern und Entscheidungsträgern unter der Überschrift der Menschzentriertheit eine umsichtige und umfassende KI-Gestaltung ermöglichen sollen. Die Realisierung der Zielsetzungen einer MZKG wirft dabei eine Vielzahl von Fragen auf. In der Dissertation habe ich mich mit zwei solcher Fragestellungen auseinandergesetzt – (1) auf konzeptioneller Ebene mit der Frage, welche Vorstellungen menschlichen Lebens und Handelns in der MZKG verwendet werden können, um den ethischen und sozialen Implikationen gerecht zu werden, und (2) auf methodischer Ebene mit der Frage, wie gestaltungstheoretische Erkenntnisse zur MZKG für die Gestaltungspraxis operationalisiert werden können. Um Antwortbeiträge zu diesen Fragen zu erarbeiten, habe ich die Rollenkonzepte und Methoden der Ansätze CTA und VSD vergleichend analysiert. Beide Ansätze sind im Feld TA gebräuchlich und in ihnen werden Denk- und Handlungsweisen angewendet und vorgeschlagen, um TG-Prozesse ganzheitlich zu betrachten und dabei besonders die auch in der MZKG bedeutsamen sozialen und ethischen Gestaltungsfaktoren in den Mittelpunkt von Gestaltungsaktivitäten zu rücken. Meine Untersuchung hat gezeigt, dass zur Realisierung von menschenzentrierten KI-Systemen auf konzeptioneller Ebene verschiedene Abstraktionsebenen benötigt werden, um verschiedene Ausprägungen menschlichen Lebens und Handelns in Gestaltungsprozessen verorten und betrachten zu können. Die Realisierungsbestrebungen erfordern flexible gedankliche Modelle, um Möglichkeiten und Folgen gestalterischen Handelns zu antizipieren und zu extrapolieren. Zur integrativen Auseinandersetzung mit den gestalterisch tätigen sowie den von der Gestaltung betroffenen Menschen muss es möglich sein, verschiedene Perspektiven zu verknüpfen, die Mensch-Technik-Beziehungen auf Micro-, Meso- und Makroebene analysieren. Diesbezüglich hat die HCAI-Community in den vergangenen Jahren festgestellt, dass sie im Anstellen von Betrachtungen auf und zwischen Meso- und Makroebenen von benachbarten Feldern lernen und Erfahrungswerte aufgreifen kann. Das Arbeiten, Denken und Argumentieren auf solch höheren Abstraktionsebenen ist in der TA-Community Tradition, weshalb sich eine Auseinandersetzung mit vor zur MZKG vergleichbaren Hintergründen entwickelten Gedanken und Ideen für die HCAI-Community lohnen kann. Im Fokus meiner Arbeit stehen deshalb die beiden Gestaltungsansätze CTA und VSD, die mehrheitlich in den 1990er- und 2000er-Jahren entwickelt wurden, um eine mensch- und werteorientierte Gestaltung technischer Systeme zu fördern. In den Rollenkonzepten und Methoden in CTA kommen mehrheitlich techniksoziologische Erkenntnisse und Perspektiven zum Tragen, die seit Mitte der 2010er-Jahre keinerlei Modernisierung mehr erfahren haben. Der Ansatz dient mehrheitlich innerhalb der TA-Community als gelegentliche Referenz, um eine Beitragsfähigkeit zu TG-Prozessen in Anspruch zu nehmen. Bei näherer Betrachtung kann zudem festgestellt werden, dass in den Denkweisen in CTA die TA-Akteure selbst (CTA agents) sich durch eine maximale

Distanzwahrung zum tatsächlichen Gestaltungsgeschehen auszeichnen, was in starkem Kontrast zu den Traditionen in HCI und HCAI sowie den gegenwärtigen Anforderungen der MZKG-Debatte steht. Anders verhält es sich in VSD, das auch in 2024 eine hohe Relevanz für die HCI-Community und darüber hinaus für verschiedene technikphilosophische Communities erfährt, die sich aktiv am MZKG-Diskurs beteiligen. VSD wurde seit seinen frühen Stunden kontinuierlich und auch außerhalb der angestammten Community weiterentwickelt. Auch werden in VSD Gestaltungsprozesse mehr auf einer Mesoebene konzipiert, was eine Zuwendung zu Micro- und Macroperspektiven tendenziell erleichtert. In Ergänzung dazu richtet sich VSD im Vergleich mit CTA deutlich expliziter an die Zielgruppe der Gestalter und Entscheidungsträger, weshalb die Rollenkonzepte und Methoden von VSD deutlich anschlussfähiger zu vielen Fragen des MZKG-Diskurses in HCAI sind.

Durch die vergleichende Analyse der Rollenkonzepte wurde deutlich, dass in beiden Ansätzen verschieden ausgeprägte Rollenkonzepte Verwendung finden, die hinsichtlich der auch in HCAI traditionellen Unterscheidung von Gestalter- und Nutzerrollen anschlussfähig sind. Insbesondere die Gedanken und Perspektiven zur Auseinandersetzung mit den Rollenkonzepten der Unbeteiligten und Betroffenen (Impactees und indirect Stakeholders), sowie der Technikgegner besitzen für den MZKG-Diskurs besondere Relevanz, da sie, auch aufgrund bereits parallel stattfindender gesellschaftlicher Entwicklungen, stärker in die Konzeption und Planung von MZKG-Prozessen Eingang finden müssen. Aufgrund der in CTA und VSD in Anspruch genommenen Gültigkeitsgrenzen wurden auch Limitierungen der Rollenkonzepte erkennbar – beide Ansätze sollen zur Gestaltung sämtlicher Technologien eingesetzt werden können, weshalb sie auf besondere Eigenschaften von und Rollen in KIGP nicht eingehen. An diesen Stellen endet die Anschlussfähigkeit beider Ansätze zum MZKG-Diskurs, da hier die HCAI-Community aktiv werden muss, um konkrete Anknüpfungspunkte zu erarbeiten und Transferleistungen zu erbringen. Hierbei gilt es für die HCAI-Community, trotz der zur Realisierung der MZKG-Ziele notwendigen Öffnung auf konzeptioneller Ebene, ihre ausgewiesenen Stärken nicht zu vernachlässigen. Dies bedeutet zusätzlich zur Adaption eines Denkens in weitreichenden Zeitfenstern und Ursache-Wirkungs-Ketten von Mensch-Technik-Beziehungen die Möglichkeit zu wahren, den Menschen in Gestaltungsprozessen auch als das zu sehen was er ist – ein Lebewesen mit individuellen Interpretationen, Bedürfnissen, Stärken und Schwächen.

Die Untersuchung der Methoden von CTA und VSD hat deutlich werden lassen, dass in beiden Ansätzen Variationen und Adaptionen bewährter qualitativer und quantitativer Methoden vorgenommen werden. Integriert in die entwickelten zyklischen Prozessmodelle zur MZKG können sie sich als wertvolle Werkzeuge im „Werkzeugkasten“ der HCAI-Community erweisen, um die in KIGP oftmals schwer greifbaren sozialen und ethischen Aspekte zu identifizieren und in konkreten Gestaltungsprojekten und Nutzungskontexten zu explorieren und zu adressieren. Dabei sollte jedoch darauf geachtet werden, Reflexions-tätigkeiten nicht zum Selbstzweck durchzuführen, sondern die daraus resultierenden Erkenntnisse aktiv in die Gestaltungsprozesse zurückzuführen. Dies ist ein zentraler Kritik-punkt der KI-Experten, die im Rahmen einer Expertenevaluation die CTA- und VSD-Methoden unter Kriterien der Gestaltungspraxis beurteilten. Die Expertenevaluation hat zudem gezeigt, dass sich Gestalter und Entscheidungsträger in KIGP ihrer sozialen und

ethischen Verantwortung bewusst sind und dieser durch verschiedene Maßnahmen versuchen gerecht zu werden. Aus diesem Grund wissen die KI-Experten um die Vorzüge der partizipativen Elemente der CTA- und VSD-Methoden, die sie als besonders hilfreich einschätzen, um die für KIGP relevanten Gruppen und Individuen der Impactees und indirect Stakeholders zu identifizieren und zu berücksichtigen. Trotz dieser Erfahrungswerte geben die KI-Experten jedoch auch zu bedenken, dass Maßnahmen mit Stakeholderinvolvement zu den inhaltlichen Zielen von Gestaltungsprozessen beitragen müssen. So können die Vorschläge der CTA- und VSD-Methoden grundsätzlich Bestandteil von künftigen Gestaltungsprozessen der MZKG sein, jedoch müssen zuvor diverse Konkretisierungs- und Modernisierungsmaßnahmen durchgeführt werden, um den konkreten Bedarfen von Gestaltern und Entscheidungsträgern in KIGP gerecht werden zu können. Dies betrifft primär die Einbettung der Methoden in den in der Praxis etablierten zyklischen und agilen Arbeitsweisen sowie der Verknüpfung mit den in beiden Ansätzen stark vernachlässigten Tätigkeiten des Projektmanagements.

In Summe kann die HCAI-Community durch eine detaillierte Auseinandersetzung mit den Ansätzen CTA und VSD viel konzeptionelles und methodisches Erfahrungswissen der TA-Community für den MZKG-Diskurs aufgreifen. Jedoch sollte dies wenn, dann in dem Bewusstsein erfolgen, dass in der TA-Community ein fundamental anderes Verständnis von TG vorhanden ist. Da in der TA-Community primär Expertisen der Geistes- und Sozialwissenschaften vertreten sind, teilt sie selten den Wunsch und die Freude am Gestalten realer Artefakte, die im Vergleich dazu in Disziplinen wie der HCAI deutlich wahrnehmbarer gelebt und praktiziert wird. Durch die Auseinandersetzung mit den primär negativen Folgen von Technik hat die TA-Community häufig zur Aufgabe zu mahnen und zu schützen, wodurch in ihren Betrachtungsweisen die potenziell positiven Folgen von TG häufig in den Hintergrund rücken. Dies kann gegenüber den technikschaaffenden Disziplinen zu grundlegend pessimistischen bis hin zu bevormundenden Mindsets und Positionen führen, die in Kontexten zur produktiven Zusammenarbeit wie jenen in HCAI nicht selten ein Hindernis darstellen. Der produktive Umgang mit diesen unter diesem Aspekt sehr heterogenen Forschungs- und Wissenskulturen ist deshalb ausschlaggebend für den Erfolg der Zusammenarbeit beider Communities. Mit Blick auf die Verantwortungsgrenzen, die die HCAI-Community derzeit im Hinblick auf die MZKG mit Feldern wie der Ethical AI aushandelt, ist gegenwärtig nur schwer abzusehen, wie viel Verantwortung die HCAI-Community zu übernehmen bereit und fachlich im Stande ist. Nach meiner Einschätzung ist durchaus vorstellbar, dass die HCAI-Community künftig, ähnlich wie die TA-Community, die Funktion einer „Dach-Community“ übernimmt, in der konkrete Diskurse zu verschiedenen Inhalten und Detailgraden zur KI-Gestaltung abgehalten werden. Auch unter diesem Aspekt kann es sich deshalb für die HCAI-Community lohnen, künftig häufiger in den Austausch mit der TA-Community zu treten und den interdisziplinären Dialog zu fördern.

## 8 Literaturverzeichnis

- Abbate, J. (2003). *Inventing the Internet* (1st MIT pbk. ed., 2000). Braille Jymico Inc.
- Abdullah, M., Madain, A., & Jararweh, Y. (2022). ChatGPT: Fundamentals, Applications and Social Impacts. *2022 Ninth International Conference on Social Networks Analysis, Management and Security*. <https://doi.org/10.1109/snams58071.2022.10062688>
- Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. In W. S. Bainbridge (Hrsg.), *Berkshire Encyclopedia Of Human-computer Interaction* (S. 445–456). Sage Publications.
- ACM SIGCHI (Regisseur). (2016). *CHI 2010 Lifetime Research Award: Lucy Suchman* [Video recording]. [https://www.youtube.com/watch?v=nwHxMWtP\\_-M](https://www.youtube.com/watch?v=nwHxMWtP_-M)
- Acuña, S. T., de Antonio, A., Ferré, X., Maté, L., & López, M. (2001). THE SOFTWARE PROCESS: MODELING, EVALUATION AND IMPROVEMENT. In S. K. Chang (Hrsg.), *Handbook of Software Engineering & Knowledge Engineering* (S. 193–237). World Scientific.
- Adelson, B. (2003). Bringing considerations of situated action to bear on the paradigm of cognitive modeling: The 2002 Benjamin Franklin Medal in computer and cognitive science presented to Lucy Suchman. *Journal of the Franklin Institute*, 340(3), 283–292. [https://doi.org/10.1016/S0016-0032\(03\)00046-2](https://doi.org/10.1016/S0016-0032(03)00046-2)
- Adenowo, A. A., & Adenowo, B. A. (2013). Software engineering methodologies: A review of the waterfall model and object-oriented approach. *International Journal of Scientific & Engineering Research*, 4(7), 427–434.
- Ahmad, K., Bano, M., Abdelrazek, M., Arora, C., & Grundy, J. (2021). What's up with Requirements Engineering for Artificial Intelligence Systems? *2021 IEEE 29th International Requirements Engineering Conference*, 1–12. <https://doi.org/10.1109/RE51729.2021.00008>
- Akrich, M. (2023). Actor Network Theory, Bruno Latour, and the CSI. *Social Studies of Science*, 53(2), 169–173. <https://doi.org/10.1177/03063127231158102>
- Albrecht, S. (2023). *ChatGPT und andere Computermodelle zur Sprachverarbeitung – Grundlagen, Anwendungspotenziale und mögliche Auswirkungen*. <https://doi.org/10.5445/IR/1000158070>
- Aldewereld, H., & Mioch, T. (2021). Values in Design Methodologies for AI. In A. Polyvyanyy & S. Rinderle-Ma (Hrsg.), *Advanced Information Systems Engineering Workshops* (Bd. 423, S. 139–150). Springer International Publishing. [https://doi.org/10.1007/978-3-030-79022-6\\_12](https://doi.org/10.1007/978-3-030-79022-6_12)
- Amershi, S., Weld, D., Vorvoreanu, M., Fournay, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for Human-AI Interaction. In S. Brewster (Hrsg.), *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (S. 1–13). Association for Computing Machinery. <https://doi.org/10.1145/3290605.3300233>
- Asikis, T., Klinglmayr, J., Helbing, D., & Pournaras, E. (2021). How value-sensitive design can empower sustainable consumption. *Royal Society Open Science*, 8(1), 201418. <https://doi.org/10.1098/rsos.201418>



- Auernhammer, J. (2020). Human-centered AI: The role of Human-centered Design Research in the development of AI. *DRS2020: Synergy*. <https://doi.org/10.21606/drs.2020.282>
- Azenkot, S., Prasain, S., Borning, A., Fortuna, E., Ladner, R. E., & Wobbrock, J. (2011). *Enhancing independence and safety for blind and deaf-blind public transit riders*. <https://doi.org/10.1145/1978942.1979424>
- Bannon, L. J., & Ehn, P. (2012). Design: Design matters in Participatory Design. In *Routledge International Handbook of Participatory Design* (S. 37–63). Routledge.
- Baxter, K., Courage, C., & Caine, K. (2015). *Understanding Your Users: A Practical Guide to User Research Methods* (2. Aufl.). Morgan Kaufmann.
- Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., & Jeffries, R. (2001). *The agile manifesto*. <https://www.agilealliance.org/wp-content/uploads/2019/09/agile-manifesto-download-2019.pdf>
- Beekman. (2004). *Description of Ethical Bio-Technology Assessment Tools for Agriculture and Food Production. Interim Report Ethical Bio-TA Tools* (LEI). LEI. <https://library.wur.nl/webquery/wurpubs/reports/356097>
- Bélisle-Pipon, J.-C., Monteferrante, E., Roy, M.-C., & Couture, V. (2023). Artificial intelligence ethics has a black box problem. *AI & SOCIETY*, 38(4), 1507–1522. <https://doi.org/10.1007/s00146-021-01380-0>
- Bendel, O. (2019). *400 Keywords Informationsethik: Grundwissen aus Computer-, Netz- und Neue-Medien-Ethik sowie Maschinenethik*. Springer-Verlag.
- Bendoukha, L. (2009). *Ein ontologisches Rahmenwerk für Kooperationsunterstützung bei evolutionärer und partizipativer Systementwicklung* [Dissertation, Staats- und Universitätsbibliothek Hamburg Carl von Ossietzky]. <https://ediss.sub.uni-hamburg.de/handle/ediss/3092>
- Berberich, N. (2019). Künstliche Intelligenz und Ethik: KI oder nicht KI? Das ist hier nicht die Frage. In K. Kersting, C. Lampert, & C. Rothkopf (Hrsg.), *Wie Maschinen lernen* (S. 229–239). Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-26763-6\\_29](https://doi.org/10.1007/978-3-658-26763-6_29)
- Bergmann, J. R. (2004). 3.2 Ethnomethodology. In *A Companion to QUALITATIVE RESEARCH* (S. 72). SAGE Publications Ltd.
- Bevan, N., Carter, J., & Harker, S. (2015). ISO 9241-11 Revised: What Have We Learnt About Usability Since 1998? In M. Kurosu (Hrsg.), *Human-Computer Interaction: Design and Evaluation* (Bd. 9169, S. 143–151). Springer International Publishing. [https://doi.org/10.1007/978-3-319-20901-2\\_13](https://doi.org/10.1007/978-3-319-20901-2_13)
- Bhasker, N., Schön, F., Kühn, J. P., & Speidel, S. (2023). Bildgebende Diagnostik und der Einsatz von künstlicher Intelligenz beim Management von Organmetastasen. *Die Onkologie*, 29(3), 182–191. <https://doi.org/10.1007/s00761-022-01282-w>

- Bijker, W. E., Hughes, T. P., & Pinch, T. J. (Hrsg.). (1987). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. MIT Press.
- Bingley, W. J., Curtis, C., Lockey, S., Bialkowski, A., Gillespie, N., Haslam, S. A., Ko, R. K. L., Steffens, N., Wiles, J., & Worthy, P. (2023). Where is the human in human-centered AI? Insights from developer priorities and user experiences. *Computers in Human Behavior*, 141, 107617. <https://doi.org/10.1016/j.chb.2022.107617>
- Björgvinsson, E., Ehn, P., & Hillgren, P.-A. (2010). Participatory design and „democratizing innovation“. *Proceedings of the 11th Biennial Participatory Design Conference*, 41–50. <https://doi.org/10.1145/1900441.1900448>
- Blomberg, J. L., & Henderson, A. (1990). Reflections on participatory design: Lessons from the trillium experience. *Proceedings of the SIGCHI conference on Human factors in computing systems Empowering people - CHI '90*. <https://doi.org/10.1145/97243.97307>
- Boehm, B. (2006). A view of 20th and 21st century software engineering. *Proceedings of the 28th International Conference on Software Engineering*, 12–29. <https://doi.org/10.1145/1134285.1134288>
- Borghoff, U. M., & Schlichter, J. H. (2000). Computer-Supported Cooperative Work. In U. M. Borghoff & J. H. Schlichter, *Computer-Supported Cooperative Work* (S. 87–141). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-662-04232-8\\_2](https://doi.org/10.1007/978-3-662-04232-8_2)
- Borning, A., & Muller, M. (2012). Next steps for value sensitive design. In J. A. Konstan, E. H. Chi, & K. Höök (Hrsg.), *CHI 2012* (S. 1125–1134). Association for Computing Machinery. <https://doi.org/10.1145/2207676.2208560>
- Böschen, S., Büscher, C., & Lösch, A. (2018). Die Forcierung des sozio-technischen Wandels. Neue (alte?) Herausforderungen für die Technikfolgenabschätzung (TA). In M. Decker, R. Lindner, S. Lingner, C. Scherz, & M. Sotoudeh (Hrsg.), *„Grand Challenges“ meistern* (S. 85–96). Nomos. <https://doi.org/10.5771/9783845283562-85>
- Boy, G. A. (Hrsg.). (2011). *The handbook of human-machine interaction: A human-centered design approach* (Online-Ausg). Ashgate.
- Bratteteig, T., Bødker, K., Dittrich, Y., Mogensen, P. H., & Simonsen, J. (2012). Methods: Organising principles and general guidelines for participatory design projects. In J. Simonsen & T. Robertson (Hrsg.), *Routledge International Handbook of Participatory Design* (S. 117–144). Routledge.
- Braun, R., & Starkbaum, J. (2023). Stakeholders in Research and Innovation: Towards Responsible Governance. In *PUTTING RESPONSIBLE RESEARCH AND INNOVATION INTO PRACTICE* (S. 229–247). SPRINGER INTERNATIONAL PU. [https://doi.org/10.1007/978-3-031-14710-4\\_12](https://doi.org/10.1007/978-3-031-14710-4_12)
- Breitsameter, C. (2015). Die Semantik „moralischer Gefühle“ zwischen Aktion, Reaktion und Interaktion. *Münchener Theologische Zeitschrift*, 66(3), 243–256.
- Brey, P. A. E. (2010). Values in Technology and Disclosive Computer Ethics. *The Cambridge Handbook of Information and Computer Ethics*, 41–58.
- Brown, T. (2008). Design thinking. *Harvard business review*, 86(6), 84.

- Bühler, P., Schlaich, P., & Sinner, D. (2017). Gestaltgesetze. In P. Bühler, P. Schlaich, & D. Sinner, *Visuelle Kommunikation* (S. 29–38). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-662-53770-1\\_3](https://doi.org/10.1007/978-3-662-53770-1_3)
- Burget, M., Bardone, E., & Pedaste, M. (2017). Definitions and Conceptual Dimensions of Responsible Research and Innovation: A Literature Review. *Science and Engineering Ethics*, 23(1), 1–19. <https://doi.org/10.1007/s11948-016-9782-1>
- Callon, M. (1984). Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay. *The Sociological Review*, 32(1\_suppl), 196–233. <https://doi.org/10.1111/j.1467-954X.1984.tb00113.x>
- Callon, M., Law, J., & Rip, A. (Hrsg.). (1986). *Mapping the Dynamics of Science and Technology*. Palgrave Macmillan UK. <https://doi.org/10.1007/978-1-349-07408-2>
- Cao, L. (2023). AI in Finance: Challenges, Techniques, and Opportunities. *ACM Computing Surveys*, 55(3), 1–38. <https://doi.org/10.1145/3502289>
- Capel, T., & Brereton, M. (2023). What is Human-Centered about Human-Centered AI? A Map of the Research Landscape. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3544548.3580959>
- Capurro, R. (2004). Informationsethik—Eine Standortbestimmung. *International Journal of Information Ethics*, 1(2004), 4–10.
- Carroll, J. M. (1997). HUMAN-COMPUTER INTERACTION: Psychology as a Science of Design. *Annual Review of Psychology*, 48(1), 61–83. <https://doi.org/10.1146/annurev.psych.48.1.61>
- Chen, H., Gomez, C., Huang, C.-M., & Unberath, M. (2022). Explainable medical imaging AI needs human-centered design: Guidelines and evidence from a systematic review. *NPJ digital medicine*, 5(1), 156.
- Chen, P. (1976). The entity-relationship model—Toward a unified view of data. *ACM Transactions on Database Systems*, 1(1), 9–36. <https://doi.org/10.1145/320434.320440>
- Chiou, L.-Y., Hung, P.-K., Liang, R.-H., & Wang, C.-T. (2023). Designing with AI: An Exploration of Co-Ideation with Image Generators. *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. <https://doi.org/10.1145/3563657.3596001>
- Chiu, T. K. F. (2023). The impact of Generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*, 1–17. <https://doi.org/10.1080/10494820.2023.2253861>
- Cho, S.-B., & Lee, J.-Y. (2002). A human-oriented image retrieval system using interactive genetic algorithm. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 32(3), 452–458.
- Clark, D., & Chamorro-Premuzic, T. (2024). Künstliche Intelligenz: Wie Menschen sich unentbehrlich machen. *Harvard Business Manager*. <https://www.manager-magazin.de/hbm/innovation/kuenstliche-intelligenz-wie-menschen-sich-unentbehrlich-machen-a-9c01ecf9-56f1-483a-bc3e-ae11c5c88202>
- Collingridge, D. (1980). *The Social Control of Technology*. St. Martin's Press.
- Cooley, M. (2000). Human-centered design. *Information design*, 59–81.

- Cox, M. J., & Marshall, G. (2007). Effects of ICT: Do we know what we should know? *Education and Information Technologies*, 12(2), 59–70. <https://doi.org/10.1007/s10639-007-9032-x>
- Coy, W., Gorny, P., & Kopp, H. (2013). *Menschengerechte Software als Wettbewerbsfaktor: Forschungsansätze und Anwenderergebnisse aus dem Programm „Arbeit und Technik “* (Bd. 40). Springer-Verlag.
- Cronholm, S., & Göbel, H. (2022). DESIGN PRINCIPLES FOR HUMAN-CENTRED AI. *ECIS 2022 Research Papers*. [https://aisel.aisnet.org/ecis2022\\_rp/32](https://aisel.aisnet.org/ecis2022_rp/32)
- Datenethikkommission der Bundesregierung. (2019). *Gutachten der Datenethikkommission* (Bundesministerium des Innern, für Bau und Heimat, Hrsg.). [https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?\\_\\_blob=publicationFile&v=6](https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6)
- Davis, J., & Nathan, L. P. (2015). Value sensitive design: Applications, adaptations, and critiques. In *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains* (S. 11–40). Springer Netherlands. [https://doi.org/10.1007/978-94-007-6970-0\\_3](https://doi.org/10.1007/978-94-007-6970-0_3)
- Decker, M. (2001). *Interdisciplinarity in Technology Assessment: Implementation and Its Chances and Limits*. Springer Science & Business Media.
- Deibel, K. (2011). *Understanding and Supporting the Adoption of Assistive Technologies by Adults with Reading Disabilities* [Dissertation]. <https://digital.lib.washington.edu/443/researchworks/handle/1773/16349>
- Deng, W. H., Yildirim, N., Chang, M., Eslami, M., Holstein, K., & Madaio, M. (2023). Investigating Practices and Opportunities for Cross-functional Collaboration around AI Fairness in Industry Practice. *2023 ACM Conference on Fairness, Accountability, and Transparency*. <https://doi.org/10.1145/3593013.3594037>
- Denning, T., Borning, A., Friedman, B., Gill, B. T., Kohno, T., & Maisel, W. H. (2010). Patients, pacemakers, and implantable defibrillators. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/1753326.1753462>
- Dennis, A., & Martin, P. J. (2005). Symbolic interactionism and the concept of power. *The British Journal of Sociology*, 56(2), 191–213. <https://doi.org/10.1111/j.1468-4446.2005.00055.x>
- Dignum, F., & Dignum, V. (2020). How to Center AI on Humans. In A. Saffiotti, L. Serafini, & P. Lukowicz (Hrsg.), *Proceedings of the First International Workshop on New Foundations for Human-Centered AI (NeHuAI) co-located with 24th European Conference on Artificial Intelligence (ECAI 2020)* (S. 59–62).
- Dignum, V. (2017). Responsible artificial intelligence: Designing AI for human values. *ITU Journal: ICT Discoveries*, 1(Special Issue 1). [https://www.itu.int/dms\\_pub/itu-s/opb/journal/s-journal-ictf.vol1-2018-1-p01-pdf-e.pdf](https://www.itu.int/dms_pub/itu-s/opb/journal/s-journal-ictf.vol1-2018-1-p01-pdf-e.pdf)
- Dimitriadis, Y., Martínez-Maldonado, R., & Wiley, K. (2021). Human-Centered Design Principles for Actionable Learning Analytics. In T. Tsiatsos, S. Demetriadis, A. Mikropoulos, & V. Dagdilelis (Hrsg.), *Research on E-Learning and ICT in Education: Technological, Pedagogical and Instructional Perspectives* (S. 277–296). Springer International Publishing. [https://doi.org/10.1007/978-3-030-64363-8\\_15](https://doi.org/10.1007/978-3-030-64363-8_15)

DIN EN ISO 9241-210:2020-03. (2020). Beuth Verlag GmbH.  
<https://doi.org/10.31030/3104744>

DIN EN ISO 27500:2017-07. (2017). Beuth Verlag GmbH.  
<https://doi.org/10.31030/2666115>

Dittrich, Y., Floyd, C., & Klischewski, R. (Hrsg.). (2002). *Social Thinking—Software Practice*. MIT Press.

Dodier-Lazaro, S., Abu-Salma, R., Becker, I., & Sasse, M. A. (2017). From Paternalistic to User-Centred Security: Putting Users First with Value-Sensitive Design. *CHI 2017 Workshop on Values in Computing*. <https://discovery.ucl.ac.uk/id/eprint/1557895/>

Doorn, N., Schuurbiers, D., van de Poel, I., & Gorman, M. E. (2013). Early Engagement and New Technologies: Towards Comprehensive Technology Engagement? In N. Doorn, D. Schuurbiers, I. van de Poel, & M. E. Gorman (Hrsg.), *Early Engagement and New Technologies: Opening Up the Laboratory* (Bd. 16, S. 233–251). Springer Netherlands.

Dorst, K. (2011). The core of 'design thinking' and its application. *Design studies*, 32(6), 521–532.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>

Eckstein, J. (2013). *Agile software development in the large: Diving into the deep*. Addison-Wesley.

Elkind, J., Black, M. S., & Murray, C. (1996). Computer-based compensation of adult reading disabilities. *Annals of Dyslexia*, 46(1), 159–186.  
<https://doi.org/10.1007/BF02648175>

European Commission. (2019). *High-Level Expert Group on Artificial Intelligence*. 2–36.

European Parliament. (2024, März 13). *EU Artificial Intelligence Act*. [https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138-FNL-COR01\\_EN.pdf](https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138-FNL-COR01_EN.pdf)

European Parliament, Directorate-General for Parliamentary Research Services, Huijstee, M., Boheemen, P., Das, D., Nierling, L., Jahnel, J., Karaboga, M., & Fatun, M. (2021). *Tackling deepfakes in European policy*. European Parliament.  
<https://doi.org/10.2861/325063>

Fan, X., & Yen, J. (2007). R-CAST: Integrating team intelligence for human-centered teamwork. *Proceedings of the national conference on artificial intelligence*, 22(2), 1535–1541.

Feng, K. J. K., Li, T. W., & Zhang, A. X. (2023). Understanding Collaborative Practices and Tools of Professional UX Practitioners in Software Organizations. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*.  
<https://doi.org/10.1145/3544548.3581273>

Firth, D. R., & Triche, J. (2023). Generative AI in practice: A Teaching Case in the Introduction to MIS class. *Proceedings of the ISCAP Conference*, 2473, 4901–4918.

- Fischer, G. (1986). II. Menschengerechte Computersysteme—Mehr als ein Schlagwort? In G. Fischer & R. Gunzenhäuser (Hrsg.), *Methoden und Werkzeuge zur Gestaltung benutzergerechter Computersysteme* (S. 17–45). De Gruyter. <https://doi.org/10.1515/9783110886252-004>
- Fischer, G. (1995). Rethinking and reinventing Artificial Intelligence from the perspective of human-centered computational artifacts. *Advances in Artificial Intelligence*, 1–11. <https://doi.org/10.1007/BFb0034793>
- Fisher, E., O'Rourke, M., Evans, R., Kennedy, E. B., Gorman, M. E., & Seager, T. P. (2015). Mapping the integrative field: Taking stock of socio-technical collaborations. *Journal of Responsible Innovation*, 2(1), 39–61. <https://doi.org/10.1080/23299460.2014.1001671>
- Floridi, L. (2021). Establishing the Rules for Building Trustworthy AI. In L. Floridi (Hrsg.), *Ethics, Governance, and Policies in Artificial Intelligence* (Bd. 144, S. 41–45). Springer International Publishing. [https://doi.org/10.1007/978-3-030-81907-1\\_4](https://doi.org/10.1007/978-3-030-81907-1_4)
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2021). An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. In *Ethics, Governance, and Policies in Artificial Intelligence* (S. 19–39). Springer, Cham. [https://link.springer.com/chapter/10.1007/978-3-030-81907-1\\_3](https://link.springer.com/chapter/10.1007/978-3-030-81907-1_3)
- Floyd, C. (1984). A Systematic Look at Prototyping. In R. Budde (Hrsg.), *Approaches to Prototyping* (S. 1–18). Springer Berlin Heidelberg. [https://link.springer.com/chapter/10.1007/978-3-642-69796-8\\_1](https://link.springer.com/chapter/10.1007/978-3-642-69796-8_1)
- Floyd, C. (1987). Steps—Eine Orientierung der Softwaretechnik auf Sozialverträgliche Technikgestaltung. In W. Schönpflug (Hrsg.), *Software-Ergonomie '87* (S. 500–503). Teubner.
- Floyd, C. (1988). A paradigm change in software engineering. *ACM SIGSOFT Software Engineering Notes*, 13(2), 25–38. <https://doi.org/10.1145/43846.43851>
- Floyd, C. (1989). Softwareentwicklung als Realitätskonstruktion. In W.-M. Lippe (Hrsg.), *Software-Entwicklung* (S. 1–20). Springer. [https://link.springer.com/chapter/10.1007/978-3-642-74872-1\\_1](https://link.springer.com/chapter/10.1007/978-3-642-74872-1_1)
- Floyd, C. (Hrsg.). (1992). *Software development and reality construction*. Springer.
- Floyd, C. (1994). Künstliche Intelligenz—Verantwortungsvolles Handeln. In S. Krämer (Hrsg.), *Geist—Gehirn—Künstliche Intelligenz* (S. 256–278). De Gruyter.
- Floyd, C. (2005). Being critical in, on or around computing? *Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility*. <https://doi.org/10.1145/1094562.1094601>
- Floyd, C. (Regisseur). (2023). *Joseph Weizenbaum—Sein Leben, sein Werk, sein Erbe* [Video recording]. <https://www.youtube.com/watch?v=t0hUJQLUPgo>
- Floyd, C., Krabbel, A., Ratuski, S., & Wetzel, I. (1997). Zur Evolution der evolutionären Systementwicklung: Erfahrungen aus einem Krankenhausprojekt. *Informatik Spektrum*, 20(1), 13–20. <https://doi.org/10.1007/s002870050047>
- Floyd, C., Mehl, W.-M., Resin, F.-M., Schmidt, G., & Wolf, G. (1989). Out of Scandinavia: Alternative Approaches to Software Design and System Development. *Human-Computer Interaction*, 4(4), 253–350.

Floyd, C., & Piepenburg, U. (1993). STEPS — ein softwaretechnischer Projektansatz und seine arbeitswissenschaftliche Begründung. In H. Reichel (Hrsg.), *Informatik—Wirtschaft—Gesellschaft* (S. 145–150). Springer Verlag. [https://doi.org/10.1007/978-3-642-78486-6\\_20](https://doi.org/10.1007/978-3-642-78486-6_20)

Floyd, C., Reisin, F.-M., & Schmidt, G. (1989). STEPS to software development with users. In G. Goos & J. Hartmanis (Hrsg.), *ESEC '89* (S. 48–64). Springer, Berlin, Heidelberg. [https://link.springer.com/chapter/10.1007/3-540-51635-2\\_32](https://link.springer.com/chapter/10.1007/3-540-51635-2_32)

Fosso Wamba, S., Bawack, R. E., Guthrie, C., Queiroz, M. M., & Carillo, K. D. A. (2021). Are we preparing for a good AI society? A bibliometric review and research agenda. *Technological Forecasting and Social Change*, 164, 120482. <https://doi.org/10.1016/j.techfore.2020.120482>

Fougere, K. T. (1991). The future role of the systems analyst as a change agent. *Journal of Systems Management*, 42(11), 6.

Franks, D. M., & Cohen, T. (2012). Social Licence in Design: Constructive technology assessment within a mineral research and development institution. *Technological Forecasting and Social Change*, 79(7), 1229–1240. <https://doi.org/10.1016/j.techfore.2012.03.001>

Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston: Pitman. Pitman.

Friedman, B. (1996). Value-Sensitive Design. *Interactions*, 3(6), 16–23. <https://doi.org/10.1145/242485.242493>

Friedman, B. (Hrsg.). (1997). *Human Values and the Design of Computer Technology* (Bd. 72). Center for the Study of Linguistics and Information.

Friedman, B. (1999). *Value-Sensitive Design: A Research Agenda for Information Technology: Report on the May 20-21, 1999 Value-Sensitive Design Workshop*. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=6d9011f9f3f8f561aa9201f556c2b73a729015c6>

Friedman, B. (2004). Value sensitive design. In W. S. Bainbridge (Hrsg.), *Encyclopedia of Human-Computer Interaction* (S. 769–774). Berkshire Publishing Group.

Friedman, B., & Hendry, D. (2012). The envisioning cards: A toolkit for catalyzing humanistic and technical imaginations. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1145–1148. <https://doi.org/10.1145/2207676.2208562>

Friedman, B., & Hendry, D. G. (2019). *Value Sensitive Design: Shaping Technology with Moral Imagination*. The MIT Press.

Friedman, B., Hendry, D. G., & Borning, A. (2017). A Survey of Value Sensitive Design Methods. *Foundations and Trends in Human-Computer Interaction*, 11(2), 63–125. <https://doi.org/10.1561/11000000015>

Friedman, B., Kahn, P., & Borning, A. (2002). Value sensitive design: Theory and methods. *University of Washington technical report*, 2(8). <http://faculty.washington.edu/pkahn/articles/vsd-theory-methods-tr.pdf>

Friedman, B., Kahn, P., Borning, A., & Hultdtgren, A. (2013). Value Sensitive Design and Information Systems. In N. Doorn, D. Schuurbiers, I. van de Poel, & M. E. Gorman (Hrsg.), *Early Engagement and New Technologies: Opening Up the Laboratory* (S. 55–95). Springer Netherlands.

Friedman, B., & Kahn, P. H. (1992). Human agency and responsible computing: Implications for computer system design. *Journal of Systems and Software*, 17(1), 7–14. [https://doi.org/10.1016/0164-1212\(92\)90075-u](https://doi.org/10.1016/0164-1212(92)90075-u)

Friedman, B., & Kahn, P. H. (2007). Human values, ethics, and design. In A. Sears & J. A. Jacko (Hrsg.), *The Human-Computer Interaction Handbook* (S. 1177–1201). CRC Press.

Friedman, B., Kahn, P. H., & Borning, A. (2006). Value Sensitive Design and Information Systems. In P. Zhang & D. Galletta (Hrsg.), *Human-computer interaction and management information systems* (S. 348–372). M.E. Sharpe.

Friedman, B., Kahn, P., Hagman, J., Severson, R., & Gill, B. (2006). The Watcher and the Watched: Social Judgments About Privacy in a Public Place. *Human-Computer Interaction*, 21(2), 235–272.

Friedman, D., Pommerenke, K., Lukose, R., Millam, G., & Huberman, B. A. (2007). Searching for the sunk cost fallacy. *Experimental Economics*, 10(1), 79–104. <https://doi.org/10.1007/s10683-006-9134-0>

Garfinkel, H. (1967). *Studies in Ethnomethodology*.

Garud, R., & Ahlstrom, D. (1997). Technology assessment: A socio-cognitive perspective. *Journal of Engineering and Technology Management*, 14(1), 25–48. [https://doi.org/10.1016/S0923-4748\(97\)00005-2](https://doi.org/10.1016/S0923-4748(97)00005-2)

Gasson, S. (2003). Human-Centered vs. User-Centered Approaches to Information System Design. *Journal of Information Technology Theory and Application (JITTA)*, 5(2). <https://aisel.aisnet.org/jitta/vol5/iss2/5>

Ghazinoory, S., Abdi, M., & Azadegan-Mehr, M. (2011). SWOT METHODOLOGY: A STATE-OF-THE-ART REVIEW FOR THE PAST, A FRAMEWORK FOR THE FUTURE. *Journal of Business Economics and Management*, 12(1), 24–48. <https://doi.org/10.3846/16111699.2011.555358>

Gloede, F. (2007). Unfolgsame Folgen: Begründungen und Implikationen der Fokussierung auf Nebenfolgen bei TA. *TATuP-Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis*, 16(1), 45–54.

Gonzatto, R. F., & van Amstel, F. M. C. (2022). User oppression in human-computer interaction: A dialectical-existential perspective. *Aslib Journal of Information Management*, 74(5), 758–781. <https://doi.org/10.1108/AJIM-08-2021-0233>

Green, K., Hull, R., McMeekin, A., & Walsh, V. (1999). The construction of the techno-economic: Networks vs. paradigms. *Research Policy*, 28(7), 777–792.

Gross, T., & Koch, M. (2007). *Computer-Supported Cooperative Work* (M. Herczeg, Hrsg.). Oldenbourg Wissenschaftsverlag GmbH. <https://doi.org/10.1524/9783486593419>

Grudin, J. (1994). Computer-supported cooperative work: History and focus. *Computer*, 27(5), 19–26. <https://doi.org/10.1109/2.291294>

Grunwald, A. (2003). Technikgestaltung – eine Einführung in die Thematik. In *Technikgestaltung zwischen Wunsch und Wirklichkeit* (S. 1–16). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-55473-5\\_1](https://doi.org/10.1007/978-3-642-55473-5_1)



- Grunwald, A. (2012). *Responsible Nanobiotechnology: Philosophy and Ethics*. CRC Press.
- Grunwald, A. (2015). Technology Assessment and Design for Values. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Hrsg.), *Handbook of Ethics, Values, and Technological Design. Sources, Theory, Values and Application Domains* (S. 67–86). Springer. <https://doi.org/10.1007/978-94-007-6970-0>
- Grunwald, A. (2019). *Technology assessment in practice and theory*. Routledge.
- Grunwald, A. (2020). Ethik und Technik. In M. Heßler & K. Liggieri (Hrsg.), *Technikanthropologie* (S. 69–82). Nomos. <https://doi.org/10.5771/9783845287959-69>
- Grunwald, A. (2021). Technische Zukunft des Menschen?: Eschatologische Erzählungen zur Digitalisierung und ihre Kritik. In O. Mitscherlich-Schönherr (Hrsg.), *Das Gelingen der künstlichen Natürlichkeit* (S. 313–332). De Gruyter. <https://doi.org/10.1515/9783110756432-015>
- Gurl, E. (2017). *SWOT analysis: A theoretical review*. <https://demo.dspacedirect.org/handle/10673/792>
- Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, 30(1), 99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Haita-Falah, C. (2017). Sunk-cost fallacy and cognitive ability in individual decision-making. *Journal of Economic Psychology*, 58, 44–59.
- Harfouche, A. L., Jacobson, D. A., Kainer, D., Romero, J. C., Harfouche, A. H., Scarascia Mugnozza, G., Moshelion, M., Tuskan, G. A., Keurentjes, J. J. B., & Altman, A. (2019). Accelerating Climate Resilient Plant Breeding by Applying Next-Generation Artificial Intelligence. *Trends in Biotechnology*, 37(11), 1217–1235. <https://doi.org/10.1016/j.tibtech.2019.05.007>
- Hartikainen, M., Väänänen, K., Lehtiö, A., Ala-Luopa, S., & Olsson, T. (2022). Human-Centered AI Design in Reality: A Study of Developer Companies' Practices. *NordiCHI 2022*, 1–11. <https://doi.org/10.1145/3546155.3546677>
- Hassenzahl, M. (2013). User experience and experience design. *The encyclopedia of human-computer interaction*, 2, 1–14.
- Hassenzahl, M., & Tractinsky, N. (2006). User experience—A research agenda. *Behaviour & Information Technology*, 25(2), 91–97. <https://doi.org/10.1080/01449290500330331>
- Heesen, J. (2021). Informationsethik. In A. Grunwald & R. Hillerbrand (Hrsg.), *Handbuch Technikethik* (S. 219–223). J.B. Metzler Verlag, ein Teil von Springer Nature.
- Heil, R. (2021). Künstliche Intelligenz/Maschinelles Lernen. In A. Grunwald & R. Hillerbrand (Hrsg.), *Handbuch Technikethik* (S. 424–428). J.B. Metzler. [https://doi.org/10.1007/978-3-476-04901-8\\_81](https://doi.org/10.1007/978-3-476-04901-8_81)
- Heinecke, A. M. (2012). Menschzentrierte Systementwicklung. In A. M. Heinecke, *Mensch-Computer-Interaktion* (S. 337–369). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-13507-1\\_14](https://doi.org/10.1007/978-3-642-13507-1_14)

- Heinlein, M., & Huchler, N. (Hrsg.). (2024). *Künstliche Intelligenz, Mensch und Gesellschaft: Soziale Dynamiken und gesellschaftliche Folgen einer technologischen Innovation*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-43521-9>
- Heiskanen, E. (2005). Taming the Golem—An Experiment in Participatory and Constructive Technology Assessment. *Science & Technology Studies*, 18(1), Article 1. <https://doi.org/10.23987/sts.55187>
- Helander, M. G. (2014). *Handbook of Human-Computer Interaction*. Elsevier.
- Hellige, H. D. (1995). *Technikgestaltung: Ein Begriff als Programm: Geschichte, Systematik und Probleme* (artec-Paper 40, S. 32). Universität Bremen. <https://www.ssoar.info/ssoar/handle/document/22008#%0Ahttps://nbn-resolving.org/urn:nbn:de:0168-ssoar-220084>
- Helms, M. M., & Nixon, J. (2010). Exploring SWOT analysis—where are we now? A review of academic research from the last decade. *Journal of strategy and management*, 3(3), 215–251.
- Hendry, D., & Friedman, B. (2023). *Resilience Grammar: A Value Sensitive Design Method for Resilience Thinking*. <https://digitalcommons.law.uw.edu/techlab/21/>
- Hernández-Ramírez, R. (2020). *The Meaning of 'Good Design' in the Age of Smart Automation: 100-114 Páginas / Journal of Science and Technology of the Arts*, v. 12 n. 3 (2020): xCoAx 2020. <https://doi.org/10.34632/JSTA.2020.9490>
- Herrmann, T., & Pfeiffer, S. (2023). Keeping the organization in the loop: A socio-technical extension of human-centered artificial intelligence. *AI & SOCIETY*, 38(4), 1523–1542. <https://doi.org/10.1007/s00146-022-01391-5>
- Heyn, H.-M., Knauss, E., Muhammad, A. P., Eriksson, O., Linder, J., Subbiah, P., Pradhan, S. K., & Tungal, S. (2021). *Requirement Engineering Challenges for AI-intense Systems Development* (arXiv:2103.10270). arXiv. <https://doi.org/10.48550/arXiv.2103.10270>
- Høstgaard, A. M. B., Bertelsen, P., & Nøhr, C. (2017). Constructive eHealth evaluation: Lessons from evaluation of EHR development in 4 Danish hospitals. *BMC medical informatics and decision making*, 17(1), 45. <https://doi.org/10.1186/s12911-017-0444-2>
- Hron, M., & Obwegeser, N. (2022). Why and how is Scrum being adapted in practice: A systematic review. *Journal of Systems and Software*, 183, 111110.
- Hronszky, I. (2005). in Future Society-Technology Relations. Remarks on Prospective Technology Analysis. In G. Banse, I. Hronszky, & G. L. Nelson (Hrsg.), *Rationality in an Uncertain World* (S. 38–59). Nomos.
- Hubig, C., Huning, A., & Ropohl, G. (Hrsg.). (2013). *Nachdenken über Technik: Die Klassiker der Technikphilosophie und neuere Entwicklungen* (3., neu bearb. und erw. Aufl., Darmstädter Ausg). edition sigma.
- Humane AI Net. (2024). *Humane AI | Human-Centered Artificial Intelligence*. <https://www.humane-ai.eu/>
- Humm, B. G., Lingner, S., Schmidt, J. C., & Wendland, K. (2021). KI-Systeme: Aktuelle Trends und Entwicklungen aus Perspektive der Technikfolgenabschätzung. *TATuP - Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis*, 30(3), Article 3. <https://doi.org/10.14512/tatup.30.3.11>
- Hybrid Human Artificial Intelligence (HHAi). (2023). *HHAi2023*. <https://hhai-conference.org/2023/>

- Hyysalo, S., & Johnson, M. (2015). The user as relational entity. *Information Technology & People*, 28(1), 72–89. <https://doi.org/10.1108/ITP-01-2014-0011>
- Inkpen, K., Chancellor, S., de Choudhury, M., Veale, M., & Baumer, E. P. S. (2019). Where is the Human? *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3290607.3299002>
- Institute for Experiential AI (Regisseur). (2023). „What is Human-Centered AI?“—Valuable Insights from Peter Norvig from Google and Stanford HAI [Video recording]. [https://www.youtube.com/watch?v=G5BRhZIAI\\_Q](https://www.youtube.com/watch?v=G5BRhZIAI_Q)
- Interaction Design Foundation. (2024, Januar 11). *What Is Human-Centered AI (HCAI)?* The Interaction Design Foundation. <https://www.interaction-design.org/literature/topics/human-centered-ai>
- Jacobs, M., Kurtz, C., Simon, J., & Böhmman, T. (2021). Value Sensitive Design and power in socio-technical ecosystems. *Internet Policy Review*, 10(3), 1–26. <https://doi.org/10.14763/2021.3.1580>
- Jiang, H. H., Brown, L., Cheng, J., Khan, M., Gupta, A., Workman, D., Hanna, A., Flowers, J., & Geburu, T. (2023). AI Art and its Impact on Artists. *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 363–374. <https://doi.org/10.1145/3600211.3604681>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Jokela, T., Iivari, N., Matero, J., & Karukka, M. (2003). The standard of user-centered design and the standard definition of usability: Analyzing ISO 13407 against ISO 9241-11. *Proceedings of the Latin American Conference on Human-Computer Interaction*, 53–60. <https://doi.org/10.1145/944519.944525>
- Jonas, H. (2019). *Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation* (7. Auflage, Bd. 3492). Suhrkamp.
- Karaman Aksentijević, N., Ježić, Z., & Zaninović, P. A. (2021). The effects of information and communication technology (ICT) use on human development—A macroeconomic approach. *Economies*, 9(3), 128.
- Kensing, F., & Blomberg, J. (1998). Participatory Design: Issues and Concerns. *Computer Supported Cooperative Work (CSCW)*, 7(3–4), 167–185. <https://doi.org/10.1023/A:1008689307411>
- Kepuska, V., & Bohouta, G. (2018). Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*. <https://doi.org/10.1109/ccwc.2018.8301638>
- Kim, S. Y., Schmitt, B. H., & Thalmann, N. M. (2019). Eliza in the uncanny valley: Anthropomorphizing consumer robots increases their perceived warmth but decreases liking. *Marketing Letters*, 30(1), 1–12. <https://doi.org/10.1007/s11002-019-09485-9>



- Law, J. (1992). Notes on the theory of the actor-network: Ordering, strategy, and heterogeneity. *Systems Practice*, 5(4), 379–393. <https://doi.org/10.1007/BF01059830>
- Lee, D., Lee, Y., Shin, Y., & Oakley, I. (2018). Designing Socially Acceptable Hand-to-Face Input. *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, 711–723. <https://doi.org/10.1145/3242587.3242642>
- Lee, S. (2023). Enhancing Video Storyboarding with Artificial Intelligence: An Integrated Approach Using ChatGPT and Midjourney within AiSAC. *International Journal of Advanced Culture Technology*, 11(3), 253–259.
- Lenzen, M. (2023). Informationsethik. In C. Neuhäuser, M.-L. Raters, & R. Stoecker (Hrsg.), *Handbuch Angewandte Ethik* (S. 363–369). J.B. Metzler. [https://doi.org/10.1007/978-3-476-05869-0\\_51](https://doi.org/10.1007/978-3-476-05869-0_51)
- Letch, N. (2010). *Engagement in e-Government Business Case Justification?: A Case for Constructive Technology Assessment*. <https://aisel.aisnet.org/amcis2010/499>
- Lewin, K. (1947). Frontiers in Group Dynamics: Concept, Method and Reality in Social Science; Social Equilibria and Social Change. *Human Relations*, 1(1), 5–41. <https://doi.org/10.1177/001872674700100103>
- Lewis, J. R. (2014). Usability: Lessons Learned ... and Yet to Be Learned. *International Journal of Human-Computer Interaction*, 30(9), 663–684. <https://doi.org/10.1080/10447318.2014.930311>
- Li, K., Tiwari, A., Alcock, J., & Bermell-Garcia, P. (2016). Categorisation of visualisation methods to support the design of Human-Computer Interaction Systems. *Applied ergonomics*, 55, 85–107.
- Liedtka, J. (2018). Why design thinking works. *Harvard Business Review*, 96(5), 72–79.
- Liggieri, K., & Müller, O. (Hrsg.). (2019). *Mensch-Maschine-Interaktion: Handbuch zu Geschichte – Kultur – Ethik*. J.B. Metzler. <https://doi.org/10.1007/978-3-476-05604-7>
- Lindsay, P., & Norman, D. (1972). *Human information processing: An introduction to psychology*. Academic Press.
- Linton, R. (1936). *The study of man: An introduction*.
- Linton, R. (1952). *The Cultural Background of Personality* (3.). Routledge & Kegan Pual LTD.
- Lorenzen, P. (1968). *Methodisches Denken*. Suhrkamp.
- Lorenzen, P. (1974). *Konstruktive Wissenschaftstheorie*. Suhrkamp.
- Lorenzen, P. (1987). *Constructive philosophy*. University of Massachusetts Press.
- Lorenzen, P., & Lorenz, K. (1978). *Dialogische Logik*. Wissenschaftliche Buchgesellschaft, [Abt. Verlag].
- Lunenburg, F. C. (2010). Managing change: The role of the change agent. *International journal of management, business, and administration*, 13(1), 1–6.
- Lüthi, N., Matt, C., & Myrach, T. (2021). A value-sensitive design approach to minimize value tensions in software-based risk-assessment instruments. *Journal of Decision Systems*, 30(2–3), 194–214. <https://doi.org/10.1080/12460125.2020.1859744>

- Lutz, B. (Hrsg.). (1987). *Technik und sozialer Wandel: Verhandlungen des 23. Deutschen Soziologentages in Hamburg 1986*.
- Macaulay, L. A. (2012). *Requirements engineering*. Springer Science & Business Media.
- MacKenzie, D. A., & Wajcman, J. (Hrsg.). (1999). *The social shaping of technology* (2nd edition). Open University Press.
- MacKenzie, I. S. (2024). *Human-computer interaction: An empirical research perspective*.
- Mädche, A., & Wallach, D. (2014). Agile und Nutzerzentrierte Softwareentwicklung. *Wissenschaft trifft Praxis: Usability betrieblicher IT-Anwendungen*, 1, 24–30.
- Madsen, K. H., & Aiken, P. H. (1993). Experiences using cooperative interactive storyboard prototyping. *Communications of the ACM*, 36(6), 57–64. <https://doi.org/10.1145/153571.163268>
- Magdans, F. (2024, Januar 25). *Welche Auswirkungen hat Künstliche Intelligenz?* <https://blog.vdi.de/welche-auswirkungen-hat-kuenstliche-intelligenz>
- Maher, M. L. (2022). *Keynote 2: Human-Centered Co-Creative AI: From Inspirational to Responsible AI*. NeurIPS2022. [https://slideslive.com/embed/presentation/38996845?js\\_embed\\_version=3&embed\\_init\\_to\\_ken=eyJhbGciOiJIUzI1NiJ9.eyJpYXQiOiJlOTg0NjU5LTlOTGtNzcxNThkMWNmY2I2liwiaSI6bnVsbCwiZSI6bnVsbCwiZmFsc2V9LCJkIjoibmlwcy5jYyJ9.JNSE8F\\_Rf165eS9ZBcgZ740COFIEWFSCMEkxNoGn4rk&embed\\_parent\\_url=https%3A%2F%2Fnips.cc%2Fvirtual%2F2022%2F65984&embed\\_origin=https%3A%2F%2Fnips.cc&embed\\_container\\_id=presentation-embed-38996845&auto\\_load=true&auto\\_play=false&zoom\\_ratio=&disable\\_fullscreen=false&locale=de&vertical\\_enabled=true&vertical\\_enabled\\_on\\_mobile=false&allow\\_hidden\\_controls\\_when\\_paused=true&fit\\_to\\_viewport=true&custom\\_user\\_id=&user\\_uuid=e47cb1f9-5458-4659-9e98-77158d1cfc6](https://slideslive.com/embed/presentation/38996845?js_embed_version=3&embed_init_to_ken=eyJhbGciOiJIUzI1NiJ9.eyJpYXQiOiJlOTg0NjU5LTlOTGtNzcxNThkMWNmY2I2liwiaSI6bnVsbCwiZSI6bnVsbCwiZmFsc2V9LCJkIjoibmlwcy5jYyJ9.JNSE8F_Rf165eS9ZBcgZ740COFIEWFSCMEkxNoGn4rk&embed_parent_url=https%3A%2F%2Fnips.cc%2Fvirtual%2F2022%2F65984&embed_origin=https%3A%2F%2Fnips.cc&embed_container_id=presentation-embed-38996845&auto_load=true&auto_play=false&zoom_ratio=&disable_fullscreen=false&locale=de&vertical_enabled=true&vertical_enabled_on_mobile=false&allow_hidden_controls_when_paused=true&fit_to_viewport=true&custom_user_id=&user_uuid=e47cb1f9-5458-4659-9e98-77158d1cfc6)
- Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures*, 90, 46–60. <https://doi.org/10.1016/j.futures.2017.03.006>
- Manders-Huits, N. (2011). What Values in Design? The Challenge of Incorporating Moral Values into Design. *Science and Engineering Ethics*, 17(2), 271–287. <https://doi.org/10.1007/s11948-010-9198-2>
- Mao, J.-Y., Vredenburg, K., Smith, P. W., & Carey, T. (2005). The state of user-centered design practice. *Communications of the ACM*, 48(3), 105–109. <https://doi.org/10.1145/1047671.1047677>
- Margetis, G., Ntoa, S., Antona, M., & Stephanidis, C. (2021). HUMAN-CENTERED DESIGN OF ARTIFICIAL INTELLIGENCE. In G. Salvendy (Hrsg.), *HANDBOOK OF HUMAN FACTORS AND ERGONOMICS* (S. 1085–1106). JOHN WILEY & SONS. <https://doi.org/10.1002/9781119636113.ch42>
- Markus, M. L., & Benjamin, R. I. (1996). Change agency-the next IS frontier. *Mis Quarterly*, 385–407.
- Mazarakis, A., Bernhard-Skala, C., Braun, M., & Peters, I. (2023). What is critical for human-centered AI at work? – Toward an interdisciplinary theory. *Frontiers in Artificial Intelligence*, 6. <https://doi.org/10.3389/frai.2023.1257057>

- Meyer, J. G., Urbanowicz, R. J., Martin, P. C. N., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N., Won, K. J., Gonzalez-Hernandez, G., & Moore, J. H. (2023). ChatGPT and large language models in academia: Opportunities and challenges. *BioData Mining*, 16(1), 20. <https://doi.org/10.1186/s13040-023-00339-9>
- Meyer, M. W., & Norman, D. (2020). Changing Design Education for the 21st Century. *She Ji: The Journal of Design, Economics, and Innovation*, 6(1), 13–49. <https://doi.org/10.1016/j.sheji.2019.12.002>
- Miller, J. K., Friedman, B., & Jancke, G. (2007). Value tensions in design. *Proceedings of the 2007 international ACM conference on Conference on supporting group work - GROUP '07*. <https://doi.org/10.1145/1316624.1316668>
- Ministerie van Onderwijs en Wetenschappen. (1984). *Integratie van Wetenschap en: Technologie in de Samenleving: Beleidsnota* (18 421, Nummern 1–2). Tweede Kamer 1983-1984.
- Mitchell, R. K., Agle, B. R., & Wood, D. J. (1997). Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts. *The Academy of Management Review*, 22(4), 853. <https://doi.org/10.2307/259247>
- Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501–507. <https://doi.org/10.1038/s42256-019-0114-4>
- Moretto, S., Robinson, D., & Moniz, A. (2014). *The role of endogenous and exogenous FTA in the European high-speed railway innovation system: CTA AS THE NEXT STEP?*
- Mou, T.-Y., Jeng, T.-S., & Chen, C.-H. (2013). From storyboard to story: Animation content development. *Educational Research and Reviews*, 8(13), 1032.
- Müller, L., Budde, M., Weibel, N., Spencer, E. A., Beigl, M., & Norman, D. (2017). Learning from failure. *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. <https://doi.org/10.1145/3123024.3124460>
- Muller, M. J. (1995). Diversity and depth in participatory design: Working with a mosaic of users and other stakeholders in the software development lifecycle. *Conference Companion on Human Factors in Computing Systems - CHI '95*, 385–386. <https://doi.org/10.1145/223355.223734>
- Muller, M., Kantosalo, A., Maher, M. L., Martin, C. P., & Walsh, G. (2024). GenAICHI 2024: Generative AI and HCI at CHI 2024. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–7. <https://doi.org/10.1145/3613905.3636294>
- Myers, B. A. (1998). A brief history of human-computer interaction technology. *Interactions*, 5(2), 44–54. <https://doi.org/10.1145/274430.274436>
- Myers, B., Hollan, J., Cruz, I., Bryson, S., Bulterman, D., Catarci, T., Citrin, W., Glinert, E., Grudin, J., & Ioannidis, Y. (1996). Strategic directions in human-computer interaction. *ACM Computing Surveys*, 28(4), 794–809. <https://doi.org/10.1145/242223.246855>
- Namatherdhala, B., Mazher, N., & Sriram, G. K. (2022). Uses of Artificial Intelligence in Autonomous Driving and V2X communication. *International Research Journal of Modernization in Engineering Technology and Science*, 4(7), 1932–1936.

- Nathan, L. P., Friedman, B., Klasnja, P., Kane, S. K., & Miller, J. K. (2008). Envisioning systemic effects on persons and society throughout interactive system design. In M. van der Schijff (Hrsg.), *Proceedings of the 7th ACM conference on Designing interactive systems* (S. 1–10). <https://doi.org/10.1145/1394445.1394446>
- Nelles, J., Brandl, C., & Mertens, A. (2018). Regelkreismodell für die menschenzentrierte Gestaltung und Evaluierung einer Mensch-Roboter-Interaktion am Beispiel eines Mensch-Roboter-Arbeitsplatzes. In Gesellschaft für Arbeitswissenschaft e.V. (Hrsg.), *ARBEIT(S).WISSEN.SCHAF(F)T - Grundlage für Management & Kompetenzentwicklung*. GfA-Press. <https://gfa2018.gesellschaft-fuer-arbeitswissenschaft.de/inhalt/B.7.11.pdf>
- Nielsen, J. (1994). *Usability Engineering*. Morgan Kaufmann.
- Norman, D. (1981). The trouble with UNIX: The user interface is horrid. *Datamation*, 27, 139–150.
- Norman, D. (1982). *Learning and memory*. W. H. Freeman.
- Norman, D. (1986). Cognitive Engineering. In D. Norman & S. W. Draper (Hrsg.), *User centered system design* (S. 31–62). Lawrence Erlbaum. <https://doi.org/10.1201/b15703-3>
- Norman, D. (1988). *The psychology of everyday things*. Basic Books.
- Norman, D. (1994). *Things that make us smart: Defending human attributes in the age of the machine*. Basic Books.
- Norman, D. (2002). Emotion & design: Attractive things work better. *Interactions*, 9(4), 36–42.
- Norman, D. (2005). Human-centered design considered harmful. *Interactions*, 12(4), 14–19. <https://doi.org/10.1145/1070960.1070976>
- Norman, D. (2006). Words Matter. Talk About People—Not Customers, Not Consumers, Not Users. *Interactions*, 13(5), 49–63. <https://doi.org/10.1145/1151314.1151340>
- Norman, D. (2013). *The Design of Everyday Things: Revised and Expanded Edition* (Revised and Expanded Edition). Basic Books.
- Norman, D. (2017). Design, Business Models, and Human-Technology Teamwork. *Research-Technology Management*, 60(1), 26–30. <https://doi.org/10.1080/08956308.2017.1255051>
- Norman, D. (2019). *The Four Fundamental Principles of Human-Centered Design and Application*. Don Norman's JND.org. <https://jnd.org/the-four-fundamental-principles-ofhuman-centered-design-and-application/>
- Norman, D. (2022, Februar 26). *Humanity-Centered versus Human-Centered Design*. Don Norman's JND.Org. <https://jnd.org/humanity-centered-versus-human-centered-design/>
- Norman, D. (2023). *Design for a better world: Meaningful, sustainable, humanity centered*. The MIT Press.
- Norman, D., & Draper, S. W. (Hrsg.). (1986). *User centered system design: New perspectives on human-computer interaction*. Lawrence Erlbaum.
- Norman, D., & Euchner, J. (2023). Design for a Better World. *Research-Technology Management*, 66(3), 11–18. <https://doi.org/10.1080/08956308.2023.2183015>
- Norman, D., & Rumelhart, D. (1981). The LNR approach to human information processing. *Cognition*, 10(1–3), 235–240. [https://doi.org/10.1016/0010-0277\(81\)90051-2](https://doi.org/10.1016/0010-0277(81)90051-2)



Notess, M., & Blevis, E. (2004). *Comparing Human-Centered Design Methods from Different Disciplines: Contextual Design and PRInCiPlES*. <https://dl.designresearchsociety.org/drs-conference-papers/drs2004/researchpapers/53/>

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., NejdI, W., Vidal, M., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., ... Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *WIREs Data Mining and Knowledge Discovery*, 10(3), e1356. <https://doi.org/10.1002/widm.1356>

Ogburn, W. F. (1922). *Social Change—With Respect to Culture and Original Nature*. B. W. Huebsch.

Opitz, M. (2009). *Organisation integrierter Dienstleistungsinnovationssysteme: Ein rollenbasiertes Rahmenkonzept*. Gabler Verlag. <https://doi.org/10.1007/978-3-8349-8322-0>

Orr, K. L., Golas, K. C., & Yao, K. (1994). Storyboard development for interactive multimedia training. *Journal of Interactive Instruction Development*, 6(3), 18–31.

Orwat, C. (with Deutschland). (2019). *Diskriminierungsrisiken durch Verwendung von Algorithmen: Eine Studie, erstellt mit einer Zuwendung der Antidiskriminierungsstelle des Bundes* (1. Auflage). Nomos.

Ottaway, R. N. (1983). The Change Agent: A Taxonomy in Relation to the Change Process. *Human Relations*, 36(4), 361–392. <https://doi.org/10.1177/001872678303600403>

Oviatt, S. (2006). Human-centered design meets cognitive load theory: Designing Interfaces that help people think. In K. Nahrstedt, M. Turk, Y. Rui, W. Klas, & K. Mayer-Patel (Hrsg.), *Proceedings of the 14th ACM international conference on Multimedia*. ACM. <https://doi.org/10.1145/1180639.1180831>

Owen, R., Bessant, J. R., & Heintz, M. (Hrsg.). (2013). *Responsible innovation*. John Wiley & Sons Inc. <https://doi.org/10.1002/9781118551424>

Ozkaya, I. (2020). What Is Really Different in Engineering AI-Enabled Systems? *IEEE Software*, 37(4), 3–6. IEEE Software. <https://doi.org/10.1109/MS.2020.2993662>

Ozmen Garibay, O., Winslow, B., Andolina, S., Antona, M., Bodenschatz, A., Coursaris, C., Falco, G., Fiore, S. M., Garibay, I., Grieman, K., Havens, J. C., Jirotko, M., Kacorri, H., Karwowski, W., Kider, J., Konstan, J., Koon, S., Lopez-Gonzalez, M., Maifeld-Carucci, I., ... Xu, W. (2023). Six Human-Centered Artificial Intelligence Grand Challenges. *International Journal of Human-Computer Interaction*, 39(3), 391–437. <https://doi.org/10.1080/10447318.2022.2153320>

Pacaux-Lemoine, M.-P., Trentesaux, D., Zambrano Rey, G., & Millot, P. (2017). Designing intelligent manufacturing systems through Human-Machine Cooperation principles: A human-centered approach. *Computers & Industrial Engineering*, 111, 581–595. <https://doi.org/10.1016/j.cie.2017.05.014>

- Pagel, S., Sophie Labonté, M., Höfer, E., Lellinger, C., & Rossner, A. (2022). Disziplinenübergreifende Teams bei der nutzerzentrierten Entwicklung von Digitalprodukten in der Medienbranche. In K. Weber & S. Reinheimer (Hrsg.), *Faktor Mensch*. [https://doi.org/10.1007/978-3-658-34524-2\\_14](https://doi.org/10.1007/978-3-658-34524-2_14)
- Palmer, T. D., & Fields, N. A. (1994). Computer supported cooperative work. *Computer*, 27(5), 15–17.
- Parandian, A. (2012). *Constructive TA of newly emerging technologies: Stimulation learning by anticipation through bridging events*.
- Parandian, A., & Rip, A. (2013). Scenarios to explore the futures of the emerging technology of organic and large area electronics. *European Journal of Futures Research*, 1(1). <https://doi.org/10.1007/s40309-013-0009-2>
- Peeters, M. M. M., van Diggelen, J., van den Bosch, K., Bronkhorst, A., Neerincx, M. A., Schraagen, J. M., & Raaijmakers, S. (2021). Hybrid collective intelligence in a human–AI society. *AI & SOCIETY*, 36(1), 217–238. <https://doi.org/10.1007/s00146-020-01005-y>
- Pentenrieder, A., & Weber, J. (2020). Lucy Suchman (geb. 1951)Annelie Pentenrieder/Jutta Weber. In M. Heßler & K. Liggieri (Hrsg.), *Technikanthropologie* (S. 215–222). Nomos Verlagsgesellschaft mbH & Co. KG. <https://doi.org/10.5771/9783845287959-215>
- Permana, P. A. G. (2015). Scrum method implementation in a software development project management. *International Journal of Advanced Computer Science and Applications*, 6(9), 198–204.
- Pesch, U. (2015). Engineers and Active Responsibility. *Science and engineering ethics*, 21(4), 925–939. <https://doi.org/10.1007/s11948-014-9571-7>
- Peuckert, R. (1992). Rolle, soziale. In B. Schäfers & J. Kopp (Hrsg.), *Grundbegriffe der Soziologie* (S. 252–256). VS Verlag für Sozialwissenschaften.
- Pinch, T. J., & Bijker, W. E. (1984). The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. *Social Studies of Science*, 14(3), 399–441. <https://doi.org/10.1177/030631284014003004>
- Pipek, V., & Wulf, V. (2009). Infrastructuring: Toward an Integrated Perspective on the Design and Use of Information Technology. *Journal of the Association for Information Systems*, 10(5), 447–473. <https://doi.org/10.17705/1jais.00195>
- Ploennigs, J., & Berger, M. (2023). Analysing the usage of AI art tools for architecture. *Proceedings of the 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference*. <https://doi.org/10.35490/EC3.2023.253>
- Pohl, K. (1996). *Requirements engineering: An overview*. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=db859f8211fd561a2763033cd2d0ad95ca4af80a>
- Povse, K. (2023). Ansätze der menschenzentrierten Gestaltung: Ein Innovationsinstrument für die Sozialwirtschaft? In F. Fischer & J. Zacher (Hrsg.), *Digitale Chancen in der häuslichen Pflege nutzen* (S. 17–25). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-662-68012-4\\_3](https://doi.org/10.1007/978-3-662-68012-4_3)
- Progress Software Corporation (Hrsg.). (2024). *Human-Centered Software Design: A State of the Marketplace*. <https://www.progress.com/resources/papers/human-centered-software-design--a-state-of-the-marketplace-report>

- Quinn, R. A. (2021). Digitale Aufklärung und Aufklärung des Digitalen: Menschen als sinnliche Wesen. *Wer bist du, Mensch?: Transformationen menschlicher Selbstverständnisse im wissenschaftlich-technischen Fortschritt*, 63.
- Quist & Vergragt. (2000). *System Innovations towards Sustainability Using Stakeholder Workshops and Scenarios*.
- Raccoon, L. B. S. (1997). Fifty years of progress in software engineering. *ACM SIGSOFT Software Engineering Notes*, 22(1), 88–104. <https://doi.org/10.1145/251759.251878>
- Rae, J. (2016). Design Value Index Exemplars Outperform the S&P 500 Index (Again) and a New Crop of Design Leaders Emerge. *Design Management Review*, 27(4), 4–11. <https://doi.org/10.1111/drev.12040>
- Rafael, S. (2023). Interactors, not Users! Towards a Neutral Interaction Design. In J. Abdelnour Nocera, M. Kristin Lárusdóttir, H. Petrie, A. Piccinno, & M. Winckler (Hrsg.), *Human-Computer Interaction—INTERACT 2023* (Bd. 14145, S. 512–516). Springer International Publishing AG.
- Razzouk, R., & Shute, V. (2012). What Is Design Thinking and Why Is It Important? *Review of Educational Research*, 82(3), 330–348. <https://doi.org/10.3102/0034654312457429>
- Reed, M. S. (2008). Stakeholder participation for environmental management: A literature review. *Biological conservation*, 141(10), 2417–2431.
- Remmen, A. (1995). Greening of the fishprocessing industry in Denmark: Constructive technology assessment. In H. Reinertsen & H. Haaland (Hrsg.), *Sustainable Fish Farming*.
- Renn, O., & Roco, M. C. (2006). Nanotechnology and the need for risk governance. *Journal of Nanoparticle Research*, 8(2), 153–191. <https://doi.org/10.1007/s11051-006-9092-7>
- Responsible AI Collaborative. (2024). *Incidents*. <https://incidentdatabase.ai/apps/incidents/>
- Riedl, M. O. (2019). Human-centered artificial intelligence and machine learning. *Human Behavior and Emerging Technologies*, 1(1), 33–36. <https://doi.org/10.1002/hbe2.117>
- Rip, A. (2001). Contributions from social studies of science and constructive technology assessment. *European Science and Technology Observatory*, 94–122.
- Rip, A. (2002). *Contributions from Social Studies of Science and Constructive Technology Assessment* (Joint Research Centre - Institute for Prospective Technological Studies & European Science and Technology Observatory, Hrsg.). [https://www.researchgate.net/publication/28552297\\_Contributions\\_from\\_Social\\_Studies\\_of\\_Science\\_and\\_Constructive\\_Technology\\_Assessment](https://www.researchgate.net/publication/28552297_Contributions_from_Social_Studies_of_Science_and_Constructive_Technology_Assessment)
- Rip, A. (2008). Nanoscience and Nanotechnologies: Bridging Gaps Through Constructive Technology Assessment. In G. H. Hadorn, H. Hoffmann-Riem, S. Biber-Klemm, W. Grossenbacher-Mansuy, D. Joye, C. Pohl, U. Wiesmann, & E. Zemp (Hrsg.), *Handbook of Transdisciplinary Research* (S. 145–157). Springer Netherlands.
- Rip, A. (2018). Constructive Technology Assessment. In *Futures of Science and Technology in Society* (S. 97–114). Springer VS, Wiesbaden. [https://doi.org/10.1007/978-3-658-21754-9\\_6](https://doi.org/10.1007/978-3-658-21754-9_6)

- Rip, A., Misa, T. J., & Schot, J. (1995). Constructive Technology Assessment: A New Paradigm for Managing Technology in Society. In A. Rip, T. J. Misa, & J. Schot (Hrsg.), *Managing Technology in Society: The Approach of Constructive Technology Assessment* (S. 1–12). Pinter Publishers.
- Rip, A., & Robinson, D. K. R. (2013). Constructive Technology Assessment and the Methodology of Insertion. In N. Doorn, D. Schuurbiers, I. van de Poel, & M. E. Gorman (Hrsg.), *Early Engagement and New Technologies: Opening Up the Laboratory* (Bd. 16, S. 37–53). Springer Netherlands.
- Rip, A., & te Kulve, H. (2008). Constructive Technology Assessment and Socio-Technical Scenarios. In E. Fisher, C. Selin, & J. M. Wetmore (Hrsg.), *The Yearbook of Nanotechnology in Society, Volume I: Presenting Futures* (Bd. 1, S. 49–70). Scholars Portal.
- Ritter, F. E., Baxter, G. D., & Churchill, E. F. (2014). Foundations for Designing User-Centered Systems. *Foundations for Designing User-Centered Systems*. <https://doi.org/10.1007/978-1-4471-5134-0>
- Roberts, L. G., & Kahn, R. E. (1972). Participating Demonstrations of a Multi-Purpose Network Linking Dissimilar Computers and Terminals. In S. Winkler (Hrsg.), *Computer Communication: Impacts and Implications* (S. 41–42).
- Robinson, D. K. R. (2009). Co-evolutionary scenarios: An application to prospecting futures of the responsible development of nanotechnology. *Technological Forecasting and Social Change*, 76(9), 1222–1239. <https://doi.org/10.1016/j.techfore.2009.07.015>
- Robinson, D. K. R. (2010). *Constructive Technology Assessment of Emerging Nanotechnologies: Experiments in Interactions*. <https://doi.org/10.3990/1.6081219>
- Roschuni, C., Goodman, E., & Agogino, A. M. (2013). Communicating actionable user research for human-centered design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 27(2), 143–154. <https://doi.org/10.1017/S0890060413000048>
- Rubenstein, A. H., & Geisler, E. (1988). The Use of Indicators and Measures of the R & D Process in Evaluating Science and Technology Programmes. In J. D. Roessner (Hrsg.), *Government Innovation Policy* (S. 185–203). Palgrave Macmillan UK. [https://doi.org/10.1007/978-1-349-08882-9\\_14](https://doi.org/10.1007/978-1-349-08882-9_14)
- Rubenstein, D. S. (2021). Acquiring ethical AI. *Fla. L. Rev.*, 73, 747.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1978). A simplest systematics for the organization of turn taking for conversation. In *Studies in the organization of conversational interaction* (S. 7–55). Elsevier.
- Sadek, M., Calvo, R. A., & Mougnot, C. (2023). Designing value-sensitive AI: a critical review and recommendations for socio-technical design processes. *AI and Ethics*, 1–19. <https://doi.org/10.1007/s43681-023-00373-7>
- Saldanha, T. J. V., John-Mariadoss, B., Wu, M. X., & Mithas, S. (2021). How Information and Communication Technology Shapes the Influence of Culture on Innovation: A Country-level Analysis. *Journal of Management Information Systems*, 38(1), 108–139. <https://doi.org/10.1080/07421222.2021.1870386>
- Saleem, K., Derhab, A., Al-Muhtadi, J., & Shahzad, B. (2015). Human-oriented design of secure Machine-to-Machine communication system for e-Healthcare society. *Computers in Human Behavior*, 51, 977–985.

Sarasin, P. (2022). Schlecht gealtert: Joseph Weizenbaums "Die Macht der Computer und die Ohnmacht der Vernunft" (1976/78). *Zeithistorische Forschungen/Studies in Contemporary History*, 19(2), 402–410.

Schepis, D., Sharon Purchase, & Ellis, N. (2019). Understanding Constructive Technology Assessments from an IMP Perspective: The Case of Autonomous Vehicles. *35th IMP Conference*, 1–8.

Schneider, W. (2008). *Ergonomische Gestaltung von Benutzungsschnittstellen: Kommentar zur Grundsatznorm DIN EN ISO 9241-110* (2., vollst. überarb. Aufl.). Beuth.

Schoenherr, J. R., Abbas, R., Michael, K., Rivas, P., & Anderson, T. D. (2023). Designing AI using a human-centered approach: Explainability and accuracy toward trustworthiness. *IEEE Transactions on Technology and Society*, 4(1), 9–23.

Schot, J. (1992). Constructive Technology Assessment and Technology Dynamics: The Case of Clean Technologies. *Science, Technology, & Human Values*, 17(1), 36–56. <https://doi.org/10.1177/016224399201700103>

Schot, J. (2001). Towards New Forms of Participatory Technology Development. *Technology Analysis & Strategic Management*, 13(1), 39–52. <https://doi.org/10.1080/09537320120040437>

Schot, J., & Rip, A. (1997). The Past and Future of Constructive Technology Assessment. *Technological Forecasting and Social Change*, 54(2–3), 251–268. [https://doi.org/10.1016/s0040-1625\(96\)00180-1](https://doi.org/10.1016/s0040-1625(96)00180-1)

Schreiber, G., & Ohly, L. (Hrsg.). (2024). *Ki:Text: Diskurse über Ki-Textgeneratoren* (1st Aufl.). De Gruyter.

Schwaber, K. (1997). SCRUM Development Process. In J. Sutherland, C. Casanave, J. Miller, P. Patel, & G. Hollowell (Hrsg.), *Business Object Design and Implementation* (S. 117–134). Springer London. [https://doi.org/10.1007/978-1-4471-0947-1\\_11](https://doi.org/10.1007/978-1-4471-0947-1_11)

Schwinger, T. (2007). *Der Rollenbegriff in sozialwissenschaftlichen Theorien: Schwieriger Umgang mit einer Metapher: Vortrag an der FH Joanneum, Graz*. [https://www.eh-darmstadt.de/fileadmin/user\\_upload/lehrende/Schwinger/Rollenbegriff.pdf](https://www.eh-darmstadt.de/fileadmin/user_upload/lehrende/Schwinger/Rollenbegriff.pdf)

Seng, L. (2019). Maschinenethik und Künstliche Intelligenz. In O. Bendel (Hrsg.), *Handbuch Maschinenethik* (S. 185–205). Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-17483-5\\_13](https://doi.org/10.1007/978-3-658-17483-5_13)

Sheppard, B., Kouyoumjian, G., Sarrazin, H., & Dore, F. (2018). *The Business Value Of Design* [McKinsey Quarterly]. McKinsey & Company. <https://www.mckinsey.com/capabilities/mckinsey-design/our-insights/the-business-value-of-design#/>

Shneiderman, B. (2020). Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-centered AI Systems. *ACM Transactions on Interactive Intelligent Systems*, 10(4), 1–31. <https://doi.org/10.1145/3419764>

Shneiderman, B. (2022). *Human-centered AI*. Oxford University Press.

Shneiderman, B., & Mengan, D. (2024). *Human-Centered AI Research Groups*. Human-Centered AI. <https://hcai.site/groups/>

- Shneiderman, B., & Rose, A. (1996). Social impact statements. In B. Liffick (Hrsg.), *CQL 96: Symposium on Computers and the Quality of Life* (S. 90–96). Association for Computing Machinery. <https://doi.org/10.1145/238339.238378>
- Shore, J., & Warden, S. (2021). *The art of agile development*. O'Reilly Media, Inc.
- Siau, K., & Wang, W. (2020). Artificial intelligence (AI) ethics: Ethics of AI and ethical AI. *Journal of Database Management (JDM)*, 31(2), 74–87.
- Silvast, A., Williams, R., Hyysalo, S., Rommetveit, K., & Raab, C. (2018). Who 'Uses' Smart Grids? The Evolving Nature of User Representations in Layered Infrastructures. *Sustainability*, 10(10), 3738. <https://doi.org/10.3390/su10103738>
- Simon, J. (2017). Value-sensitive design and responsible research and innovation. In S. O. Hansson (Hrsg.), *The ethics of technology* (S. 219–236). Rowman & Littlefield International, Ltd.
- Simonsen, P. (1997). *Promoting a Development Culture in Your Organization: Using Career Development as a Change Agent*. ERIC. <https://eric.ed.gov/?id=ED422504>
- Skinner, K., Hanning, R., Sutherland, C., Edwards-Wheesk, R., & Tsuji, L. (2012). Using a SWOT Analysis to Inform Healthy Eating and Physical Activity Strategies for a Remote First Nations Community in Canada. *American journal of health promotion : AJHP*, 26, e159-70. <https://doi.org/10.4278/ajhp.061019136>
- Smith, M. R., & Marx, L. (Hrsg.). (1994). *Does Technology Drive History? The Dilemma of Technological Determinism*. MIT Press.
- Solanki, P., Grundy, J., & Hussain, W. (2023). Operationalising ethics in artificial intelligence for healthcare: A framework for AI developers. *AI and Ethics*, 3(1), 223–240. <https://doi.org/10.1007/s43681-022-00195-z>
- Spiekermann, S. (2010). Über die Bedeutung von Menschenbildern für die Gestaltung Allgegenwärtiger Technik. In M. Bölker, M. Gutmann, & W. Hesse (Hrsg.), *Information und Menschenbild* (S. 63–77). Springer, Berlin, Heidelberg. [https://link.springer.com/chapter/10.1007/978-3-642-04742-8\\_4](https://link.springer.com/chapter/10.1007/978-3-642-04742-8_4)
- Spinuzzi, C. (2005). The Methodology of Participatory Design. *Technical Communication*, 52(2), 163–174.
- Srivastava, A., Bhardwaj, S., & Saraswat, S. (2017). SCRUM model for agile methodology. *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 864–869.
- Stahl, B. C., Andreou, A., Brey, P., Hatzakis, T., Kirichenko, A., Macnish, K., Shaelou, S. L., Patel, A., Ryan, M., & Wright, D. (2021). Artificial intelligence for human flourishing—Beyond principles for machine learning. *Journal of Business Research*, 124, 374–388.
- Stanford Institute for Human-Centered Artificial Intelligence. (2023, Dezember 15). *Home*. <https://hai.stanford.edu/>
- Steen, M. (2012). Human-centered design as a fragile encounter. *Design Issues*, 28(1), 72–80.
- Steen, M., & Nauta, J. (2020). Advantages and disadvantages of societal engagement: A case study in a research and technology organization. *Journal of Responsible Innovation*, 7(3), 598–619. <https://doi.org/10.1080/23299460.2020.1813864>

- Steinfath, H. (2001). Gefühle und Werte. *Zeitschrift für philosophische Forschung*, *H. 2*, 196–220.
- Steinfeld, K. (2023). Clever little tricks: A socio-technical history of text-to-image generative models. *International Journal of Architectural Computing*, *21*(2), 211–241. <https://doi.org/10.1177/14780771231168230>
- Stern, H., & Becker, T. (2019). Concept and evaluation of a method for the integration of human factors into human-oriented work design in cyber-physical production systems. *Sustainability*, *11*(16), 4508.
- Strahringer, S. (2016). Modelle und Metamodellierung. In T. Benker & Jürck (Hrsg.), *Geschäftsprozessorientierte Systementwicklung* (S. 11–23). Springer Fachmedien Wiesbaden.
- Strough, J., Mehta, C. M., McFall, J. P., & Schuller, K. L. (2008). Are Older Adults Less Subject to the Sunk-Cost Fallacy Than Younger Adults? *Psychological Science*, *19*(7), 650–652. <https://doi.org/10.1111/j.1467-9280.2008.02138.x>
- Subramonyam, H., Im, J., Seifert, C., & Adar, E. (2022). Solving Separation-of-Concerns Problems in Collaborative Design of Human-AI Systems through Leaky Abstractions. *CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3491102.3517537>
- Suchman, L. (1985). *PLANS AND SITUATED ACTIONS - The problem of human-machine communication*. Xerox Corporation.
- Suchman, L. (1999). *Plans and situated actions: The problem of human-machine communication*. Cambridge University Press.
- Suchman, L. (2007). *Human-machine reconfigurations: Plans and situated actions* (2. ed., reprint). Cambridge Univ. Press.
- Tatum, J. S. (1997). The Political Construction of Technology: A Call for Constructive Technology Assessment. In K. Shrader-Frechette & L. Westra (Hrsg.), *Technology and values* (S. 115–128). Rowman & Littlefield Publishers.
- Teich, I. (2020). Meilensteine der Entwicklung Künstlicher Intelligenz. *Informatik Spektrum*, *43*(4), 276–284. <https://doi.org/10.1007/s00287-020-01280-5>
- The Human-Centred AI Master's Programme and Community. (2024). *HCAIM*. HCAIM. The Human-Centred AI Master's Programme and Community. Co-Financed by the Connecting Europe Facility of the European Union Under Grant No CEF-TC-2020-1 Digital Skills 2020-EU-IA-0068. <https://humancentered-ai.eu/>
- Tiwari, R. (2023). Ethical And Societal Implications of AI and Machine Learning. *Interantional Journal Of Scientific Research In Engineering And Management*, *7*(01).
- Townsend, M., & Romme, A. G. L. (2024). The Emerging Concept of the Human-Centered Organization: A Review and Synthesis of the Literature. *Humanistic Management Journal*, *9*(1), 53–74. <https://doi.org/10.1007/s41463-024-00168-w>
- Truong, K. N., Hayes, G. R., & Abowd, G. D. (2006). Storyboarding: An empirical determination of best practices and effective guidelines. *Proceedings of the 6th Conference on Designing Interactive Systems*, 12–21. <https://doi.org/10.1145/1142405.1142410>

Trust, T., Whalen, J., & Mouza, C. (2023). Editorial: ChatGPT: Challenges, Opportunities, and Implications for Teacher Education. *Contemporary Issues in Technology and Teacher Education*, 23(1), 1–23.

Umbrello, S., & Gambelin, O. (2023). Agile as a Vehicle for Values: A Value Sensitive Design Toolkit. *Rethinking Technology and Engineering*, 45, 169–181.

Umbrello, S., & van de Poel, I. (2021). Mapping value sensitive design onto AI for social good principles. *AI and Ethics*, 1(3), 283–296. <https://doi.org/10.1007/s43681-021-00038-3>

Università degli Studi di Milano Statale. (2024). *Human-Centered Artificial Intelligence*. <https://hcai.cdl.unimi.it/en>

University of Gothenburg. (2024, Januar 24). *Human-centered Artificial Intelligence Master's Programme | University of Gothenburg*. <https://www.gu.se/en/study-gothenburg/human-centered-artificial-intelligence-masters-programme-t2hai>

van de Poel, I. (2020). Three philosophical perspectives on the relation between technology and society, and how they affect the current debate about artificial intelligence. *Human Affairs*, 30(4), 499–511. <https://doi.org/10.1515/humaff-2020-0042>

van den Hoven, J. (2013). Value Sensitive Design and Responsible Innovation. In R. Owen, J. R. Bessant, & M. Heintz (Hrsg.), *Responsible innovation* (S. 75–83). John Wiley & Sons Inc. <https://doi.org/10.1002/9781118551424.ch4>

van den Hoven, J., Vermaas, P. E., & van de Poel, I. (2015). Design for Values: An Introduction. In J. van den Hoven, P. E. Vermaas, & I. de van Poel (Hrsg.), *Handbook of ethics, values, and technological design* (S. 1–7). SpringerReference.

van Merkerk, R. (2007). Intervening in emerging nanotechnologies: A CTA of Lab-on-a-chip technology. *Netherlands Geographical Studies*, 367. <https://dspace.library.uu.nl/handle/1874/24908>

Van Pelt, A., & Hey, J. (2011). Using TRIZ and human-centered design for consumer product development. *Procedia Engineering*, 9, 688–693. <https://doi.org/10.1016/J.PROENG.2011.03.156>

van Wynsberghe, A. (2013). Designing robots for care: Care centered value-sensitive design. *Science and engineering ethics*, 19(2), 407–433. <https://doi.org/10.1007/s11948-011-9343-6>

van Wynsberghe, A., & Robbins, S. (2014). Ethicist as designer: A pragmatic approach to ethics in the lab. *Science and Engineering Ethics*, 20(4), 947–961. <https://doi.org/10.1007/s11948-013-9498-4>

Velsen, L. van, Ludden, G., & Grünloh, C. (2022). The Limitations of User-and Human-Centered Design in an eHealth Context and How to Move Beyond Them. *Journal of Medical Internet Research*, 24(10), e37341. <https://doi.org/10.2196/37341>

Verbeek, P.-P. (2006). Materializing Morality. *Science, Technology, & Human Values*, 31(3), 361–380. <https://doi.org/10.1177/0162243905285847>

Versteeg, T., Baumann, M., Weil, M., & Moniz, A. B. (2017). Exploring emerging battery technology for grid-connected energy storage with constructive technology assessment. *Technological Forecasting and Social Change*, 115, 99–110. <https://doi.org/10.1016/j.techfore.2016.09.024>



- Viégas, F. (2022). Keynote 5: Building human-centric AI systems: Thoughts on user agency, transparency and trust. NeurIPS2022.  
[https://slideslive.com/embed/presentation/38996871?js\\_embed\\_version=3&embed\\_init\\_to\\_ken=eyJhbGciOiJIUzI1NiJ9.eyJpYXQiOiE3MTgxNzYyNDAsImV4cCI6MTcxODMwNTg0MCwiLCJmYWZlbnRlbnQ3Y2lxZjktNTQ1OC00NjU5LTllOTgtNzcxNThkMWNmY2I2liwiaSI6bnVsY2I2bnVsbCwicibGl6ZmFsc2V9LCJkIjoibmlwcy5jYyJ9.IFM1FZ6UI1tYvxmeggXsOUSBIjb3VvZtB-](https://slideslive.com/embed/presentation/38996871?js_embed_version=3&embed_init_to_ken=eyJhbGciOiJIUzI1NiJ9.eyJpYXQiOiE3MTgxNzYyNDAsImV4cCI6MTcxODMwNTg0MCwiLCJmYWZlbnRlbnQ3Y2lxZjktNTQ1OC00NjU5LTllOTgtNzcxNThkMWNmY2I2liwiaSI6bnVsY2I2bnVsbCwicibGl6ZmFsc2V9LCJkIjoibmlwcy5jYyJ9.IFM1FZ6UI1tYvxmeggXsOUSBIjb3VvZtB-1OgQA2xZvc&embed_parent_url=https%3A%2F%2Fnips.cc%2Fvirtual%2F2022%2F66023&embed_origin=https%3A%2F%2Fnips.cc&embed_container_id=presentation-embed-38996871&auto_load=true&auto_play=false&zoom_ratio=&disable_fullscreen=false&locale=de&vertical_enabled=true&vertical_enabled_on_mobile=false&allow_hidden_controls_when_paused=true&fit_to_viewport=true&custom_user_id=&user_uuid=e47cb1f9-5458-4659-9e98-77158d1cfcb6)  
1OgQA2xZvc&embed\_parent\_url=https%3A%2F%2Fnips.cc%2Fvirtual%2F2022%2F66023&embed\_origin=https%3A%2F%2Fnips.cc&embed\_container\_id=presentation-embed-38996871&auto\_load=true&auto\_play=false&zoom\_ratio=&disable\_fullscreen=false&locale=de&vertical\_enabled=true&vertical\_enabled\_on\_mobile=false&allow\_hidden\_controls\_when\_paused=true&fit\_to\_viewport=true&custom\_user\_id=&user\_uuid=e47cb1f9-5458-4659-9e98-77158d1cfcb6
- von Alemann, U., Schatz, H., von Alemann, U., & Schatz, H. (1986). Sozialverträgliche Technikgestaltung. Entwurf eines politischen, wissenschaftlichen, gesellschaftlichen Programms. In *Mensch und Technik* (S. 21–49). VS Verlag für Sozialwissenschaften.  
[https://doi.org/10.1007/978-3-322-85945-7\\_1](https://doi.org/10.1007/978-3-322-85945-7_1)
- von Gleich, A. (2013). Prospektive Technikbewertung und Technikgestaltung zur Umsetzung des Vorsorgeprinzips. In Georg. Simonis (Hrsg.), *Konzepte und Verfahren der Technfolgenabschätzung* (S. 51–73). Springer Fachmedien Wiesbaden and Imprint and Springer VS. [https://link.springer.com/chapter/10.1007/978-3-658-02035-4\\_4](https://link.springer.com/chapter/10.1007/978-3-658-02035-4_4)
- Von Schomberg, R. (2013). A Vision of Responsible Research and Innovation. In R. Owen, J. Bessant, & M. Heintz (Hrsg.), *Responsible Innovation* (1. Aufl., S. 51–74). Wiley.  
<https://doi.org/10.1002/9781118551424.ch3>
- Waefler, T., & Schmid, U. (2020). Explainability is not Enough: Requirements for Human-AI-Partnership in Complex Socio-Technical Systems. In F. Matos (Hrsg.), *Proceedings of the 2nd European Conference on the Impact of Artificial Intelligence and Robotics (ECIAIR 2020)* (S. 185–194). ACPIIL. <https://doi.org/10.20378/irb-49775>
- Wasielowski, A. (2023). “Midjourney Can’t Count”: Questions of Representation and Meaning for Text-to-Image Generators. *The Interdisciplinary Journal of Image Sciences*, 37(1), 71–82.
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45.  
<https://doi.org/10.1145/365153.365168>
- Weizenbaum, J. (1967). Contextual understanding by computers. *Communications of the ACM*, 10(8), 474–480. <https://doi.org/10.1145/363534.363545>
- Weizenbaum, J. (1976). *Computer power and human reason: From judgement to calculation*. W.H. Freeman.
- Weizenbaum, J. (1991). *Kurs auf den Eisberg: Die Verantwortung des einzelnen und die Diktatur der Technik* (3. Aufl., Bd. 541). Piper.
- Wendland, K. (2021a). Demystifying Artificial Consciousness – About Attributions, Black Swans, and Suffering Machines. *Journal of AI Humanities*, 9, 137–166.  
<https://doi.org/10.46397/JAIH.9.7>
- Wendland, K. (2021b). Menschenbild ohne Menschen: Subjektkonstitution im Spiegel synthetischer Konkurrenz. In A. Grunwald (Hrsg.), *Wer bist du, Mensch?* (S. 240–259). Herder Verlag.

- Wendt, G., & Klug, F. (Hrsg.). (2001). *Computermacht und Gesellschaft: Freie Reden* (1.). Suhrkamp Verlag.
- Wennker, P. (Hrsg.). (2020). *Künstliche Intelligenz in der Praxis—Anwendung in Unternehmen und Branchen: KI wettbewerbs- und zukunftsorientiert einsetzen* (1. Auflage 2020). Springer Fachmedien Wiesbaden GmbH and Springer Gabler. <https://doi.org/10.1007/978-3-658-30480-5>
- Werbeagentur IMPACT. (2023, September 5). *Find a construction or technical job in Belgium | IMPACT*. <https://impact.be/en/international-workers>
- Weydner-Volkman, S. (2021). *Technikvertrauen. Beiträge zur Technikfolgenabschätzung jenseits von Akzeptanz und Akzeptabilität?* <https://philpapers.org/rec/WEYTBZ>
- Weyer, J. (1994). *Perspektiven der sozialwissenschaftlichen Technikfolgenabschätzung*. [http://129.217.131.68:8080/bitstream/2003/34327/1/1994\\_Weyer\\_Perspektiven\\_Technikfolgenabsch%C3%A4tzung.pdf](http://129.217.131.68:8080/bitstream/2003/34327/1/1994_Weyer_Perspektiven_Technikfolgenabsch%C3%A4tzung.pdf)
- Whittlestone, J., Nyrup, R., Alexandrova, A., Dihal, K., & Cave, S. (2019). Ethical and societal implications of algorithms, data, and artificial intelligence: A roadmap for research. *London: Nuffield Foundation*. <https://www.nuffieldfoundation.org/wp-content/uploads/2019/02/Ethical-and-Societal-Implications-of-Data-and-AI-report-Nuffield-Foundat.pdf>
- Wiarda, M. J. (2023). *Responsible Innovation for Wicked Societal Challenges: An Exploration of Strengths and Limitations* (Bd. 29). 4TU.Centre for Ethics and Technology. <https://doi.org/10.4233/uuid:ab63eb8b-a2d9-4d8a-9134-b9af2df1f62b>
- Widder, D. G., & Nafus, D. (2023). Dislocated accountabilities in the “AI supply chain”: Modularity and developers’ notions of responsibility. *Big Data & Society*, 10(1). <https://doi.org/10.1177/20539517231177620>
- Wiedermann, M. (2024, Januar 16). *Storyboard erstellen: Top 10 Storyboard-Programme 2021*. <https://filmora.wondershare.de/animated-video/top-best-storyboarding-software.html>
- Wilson, E. A. (2010). *Affect and Artificial Intelligence*. University of Washington Press.
- Windl, M., Feger, S. S., Zijlstra, L., Schmidt, A., & Wozniak, P. W. (2022). ‘It Is Not Always Discovery Time’: Four Pragmatic Approaches in Designing AI Systems. *CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3491102.3501943>
- Winograd, T., & Woods, D. D. (1997). The challenge of human-centered design. *Human-centered systems: information, interactivity, and intelligence*.
- Wittpahl, V. (Hrsg.). (2019). *Künstliche Intelligenz: Technologie | Anwendung | Gesellschaft*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-58042-4>
- Woelfer, J. P., & Hendry, D. G. (2009). Stabilizing homeless young people with information and place. *Journal of the American Society for Information Science and Technology*, 60(11), 2300–2312. <https://doi.org/10.1002/asi.21146>

- Wolff, E. (Regisseur). (2021, Juli 9). *Prof. Christiane Floyd zu „menschenzentrierter Software-Entwicklung“* [Video recording]. <https://www.youtube.com/watch?v=lygh1qKm7YU>
- Wulf, V., & Rohde, M. (1995). Towards an integrated organization and technology development. *Proceedings of the conference on Designing interactive systems processes, practices, methods, & techniques - DIS '95*. <https://doi.org/10.1145/225434.225441>
- Xu, W. (2019). Toward human-centered AI: a perspective from human-computer interaction. *Interactions*, 26(4), 42–46.
- Xu, W., Dainoff, M. J., Ge, L., & Gao, Z. (2023). Transitioning to Human Interaction with AI Systems: New Challenges and Opportunities for HCI Professionals to Enable Human-Centered AI. *International Journal of Human-Computer Interaction*, 39(3), 494–518. <https://doi.org/10.1080/10447318.2022.2041900>
- Xu, W., & Gao, Z. (2023). *Enabling Human-Centered AI: A Methodological Perspective* (arXiv:2311.06703). arXiv. <http://arxiv.org/abs/2311.06703>
- Xu, W., Gao, Z., & Dainoff, M. (2023). *An HCAI Methodological Framework: Putting It Into Action to Enable Human-Centered AI* (arXiv:2311.16027). arXiv. <http://arxiv.org/abs/2311.16027>
- Xue, L., & Pang, Z. (2022). Ethical governance of artificial intelligence: An integrated analytical framework. *Journal of Digital Economy*, 1(1), 44–52. <https://doi.org/10.1016/j.jdec.2022.08.003>
- Yampolskiy, R. V. (2019). Predicting future AI failures from historic examples. *foresight*, 21(1), 138–152. <https://doi.org/10.1108/fs-04-2018-0034>
- Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376301>
- Yildirim, N., Pushkarna, M., Goyal, N., Wattenberg, M., & Viégas, F. (2023). Investigating How Practitioners Use Human-AI Guidelines: A Case Study on the People + AI Guidebook. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3544548.3580900>
- Yoo, D., Hultgren, A., Woelfer, J. P., Hendry, D. G., & Friedman, B. (2013). A value sensitive action-reflection model. In W. E. Mackay, S. Brewster, & S. Bødker (Hrsg.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 419–428). ACM. <https://doi.org/10.1145/2470654.2470715>
- Zahra, S. A. (2008). The virtuous cycle of discovery and creation of entrepreneurial opportunities. *Strategic Entrepreneurship Journal*, 2(3), 243–257. <https://doi.org/10.1002/sej.47>
- Zarte, M., Pechmann, A., & Nunes, I. L. (2020). Principles for Human-Centered System Design in Industry 4.0 – A Systematic Literature Review. In I. L. Nunes (Hrsg.), *Advances in Human Factors and Systems Interaction* (Bd. 1207, S. 140–147). Springer International Publishing. [https://doi.org/10.1007/978-3-030-51369-6\\_19](https://doi.org/10.1007/978-3-030-51369-6_19)

Zitterbarth, W. (1991). Der Erlanger Konstruktivismus in seiner Beziehung zum Konstruktiven Realismus. *Formen des Konstruktivismus in Diskussion*. Wien: WUV-Universitätsverlag. S. 73–87.

Zuber, N., Gogoll, J., Kacianka, S., Nida-Rümelin, J., & Pretschner, A. (2024). Value-Sensitive Software Design: Ethical Deliberation in Agile Development Processes. In *INTRODUCTION TO DIGITAL HUMANISM* (S. 339–358). SPRINGER INTERNATIONAL PU.

## 9 Abbildungsverzeichnis

Abbildung 1: Forschungsdesign der Dissertation .....	9
Abbildung 2: Ein menschenzentrierter Gestaltungsprozess nach DIN EN ISO 9241-210:2020-03 (2020, S. 25) .....	11
Abbildung 3: Das Wasserfallmodell.....	12
Abbildung 4: Darstellung des skandinavischen Ansatzes zur Softwareentwicklung gegen Ende der 1980er-Jahre (Floyd, Mehl, et al., 1989, S. 265) .....	16
Abbildung 5: Prozessmodell von STEPS (Floyd, Reisin, et al., 1989, S. 58) .....	17
Abbildung 6: Christiane Floyds Position zur MZG im Jahr 2023 (Floyd, 2023) .....	18
Abbildung 7: Elemente von Interaktionstätigkeiten mit Computersystemen (Norman, 1986, S. 42) .....	20
Abbildung 8: Ein menschenzentrierter Gestaltungsprozess nach Donald Norman (2013, S. 222) .....	21
Abbildung 9: Von Lucy Suchman erstelltes Template zur kombinierten Konversations- und Systemzustandsanalyse (Suchman, 1985, S. 80).....	23
Abbildung 10: Gestaltungsprinzipien der MZG, abgeleitet aus DIN EN ISO 9241-210:2020-03.....	25
Abbildung 11: Werbeplakat mit der Aufschrift „Hey Chat GPT finish this building...“ (Werbeagentur IMPACT, 2023) .....	30
Abbildung 12: Themenübersicht der HHAI-Konferenz in 2022 (Hybrid Human Artificial Intelligence (HHAI), 2023).....	33
Abbildung 13: Themenfelder in HCAI (Capel & Brereton, 2023, S. 5) .....	34
Abbildung 14: MZKG-Modell nach Xu et al. (2023, S. 8) .....	36
Abbildung 15: MZKG-Modell nach Auernhammer (2020, S. 1328) .....	37
Abbildung 16: Sechs Erweiterungen für das MZG-Modell in DIN EN ISO 9241-210 nach Margetis et al. (2021, S. 1089) .....	38
Abbildung 17: MZKG-Modell nach Margetis et al. (2021, S. 1099) .....	39
Abbildung 18: Relevante Faktoren einer MZKG nach Shneiderman (2022, S. 28).....	40
Abbildung 19: MZKG-Modell nach Shneiderman (Shneiderman, 2022, S. 86) .....	41
Abbildung 20: Nutzerauffassungen zu Responsible AI (Maher, 2022).....	43
Abbildung 21: Spektrum von Stakeholdern in der KI-Gestaltung (Viégas, 2022) .....	43



Abbildung 22: Charakterisierung von HCAI nach Norvig (Institute for Experiential AI, 2023).....	45
Abbildung 23: TA in the making of technology als ein Element einer TA in action (Grunwald, 2019, S. 89) .....	48
Abbildung 24: Die drei Aktivitäten der CTA-Methodik.....	53
Abbildung 25: In der Technologiefeldanalyse begeben sich CTA agents in ein Technologiefeld hinein .....	54
Abbildung 26: Die Erkenntnisse der Technologiefeldanalyse werden in soziotechnische Szenarien überführt .....	55
Abbildung 27: In moderierten Workshops diskutieren Akteure die soziotechnischen Szenarien.....	56
Abbildung 28: Die tripartite methodology in VSD.....	61
Abbildung 29: Zuordnung der VSD-Methoden zu den drei Untersuchungstypen der tripartite methodology.....	64
Abbildung 30: Identifikation direkter und indirekter Stakeholder (Stakeholderanalyse).....	65
Abbildung 31: Identifikation möglicher nicht-intendierter Nutzungskontexte und Folgen (Stakeholderanalyse).....	66
Abbildung 32: Identifikation potenzieller Nutzen und Risiken für Stakeholder (Stakeholderanalyse).....	66
Abbildung 33: Befragung von Stakeholdern zu potenziellen Nutzen und Risiken (Value dams and flows) .....	67
Abbildung 34: Ermittlung von Value dams und Value flows (Value dams and flows).....	67
Abbildung 35: Durchführung eines Prototyping-Workshops (Value-oriented Prototyping) .....	68
Abbildung 36: Analyse der Prototypen (Value-oriented Prototyping) .....	68
Abbildung 37: In CTA und VSD werden beim Blick auf Gestaltungsprozesse unterschiedliche Abstraktionsebenen verwendet (Doorn et al., 2013, S. 239) .....	70
Abbildung 38: Entity-Relationship-Modell des Rollenverständnisses von Ralph Linton .....	82
Abbildung 39: Elemente eines Prozessmodells in der MSEP nach Acuña et al. (2001, S. 199) .....	83
Abbildung 40: Ein in der Softwareprozessmodellierung verwendetes Rollenverständnis .....	84
Abbildung 41: Entity-Relationship-Modell des von mir verwendeten Rollenverständnisses.....	86

Abbildung 42: Zeitlicher Zusammenhang von enactment cycles und selection cycles nach Garud und Ahlstrom (1997, S. 44) .....	134
Abbildung 43: Abstraktionsebenen, mit deren Hilfe in VSD menschliches Leben und Handeln untersucht wird (B. Friedman & Hendry, 2019, S. 21) .....	156
Abbildung 44: Verwendete Rollenkonzepte in HCAI nach Shneiderman (2022, S. 30).....	161
Abbildung 45: Beispiel eines Storyboards (Wiedermann, 2024).....	177
Abbildung 46: Layout einer Ansatzseite in den Storyboards der Interviewstudie .....	178

# 10 Tabellenverzeichnis

Tabelle 1:	Auswahl von HCAI-Forschungsgruppen und -Lehrstühlen im Jahr 2024 (Einträge teilweise aus Shneiderman & Mengan, 2024) .....	32
Tabelle 2:	Layout des Kategoriensystems der vergleichenden Analyse mit Beispieleinträgen.....	89
Tabelle 3:	Zentrale Kategorien der vergleichenden Analyse im Hinblick auf Rollenkonzepte .....	90
Tabelle 4:	Übersicht der verwendeten Rollenkonzepte in CTA und VSD .....	92
Tabelle 5:	Teilnehmerfeld der Interviewstudie .....	181
Tabelle 6:	Eine typische SWOT-Matrix (Skinner et al., 2012, S. 163).....	182



# 11 Anhang

## Storyboards (deutsch)

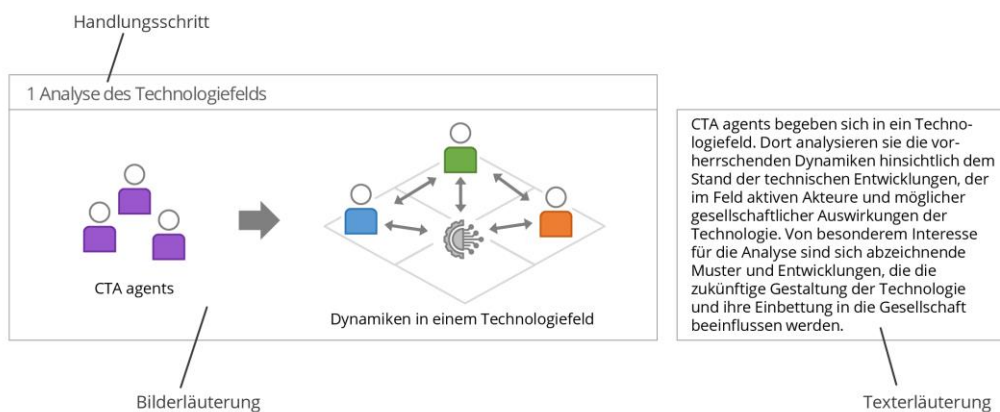


### Über dieses Dokument

Dieses Dokument ist Teil einer wissenschaftlichen Untersuchung und dient der Vorbereitung und inhaltlichen Unterstützung von Experteninterviews. Die Informationen in diesem Dokument wenden sich primär an Experten und Professionals aus dem Feld der Künstlichen Intelligenz (KI), die Projekterfahrung in der Gestaltung und Entwicklung von KI-Systemen besitzen.

In diesem Dokument werden Gestaltungsmethoden der Ansätze Constructive Technology Assessment (CTA) und Value Sensitive Design (VSD) dargestellt und erläutert. Zur Darstellung der Methoden und ihrer einzelnen Handlungsschritte wird eine Kombination aus Bild- und Textinformationen verwendet (siehe Beispiel unten). Sämtliche dafür verwandten Informationen basieren auf wissenschaftlicher Literatur.

### Beispiel für die Darstellung von Informationen im Dokument



Bei Fragen oder Anmerkungen zu diesem Dokument wenden Sie sich bitte an:

Karlsruher Institut für Technologie (KIT)  
Institut für Technikfolgenabschätzung und Systemanalyse (ITAS)

Pascal Vetter, B.Eng., M.Sc.  
Wissenschaftlicher Mitarbeiter

Karlstrasse 11  
76133 Karlsruhe, Germany

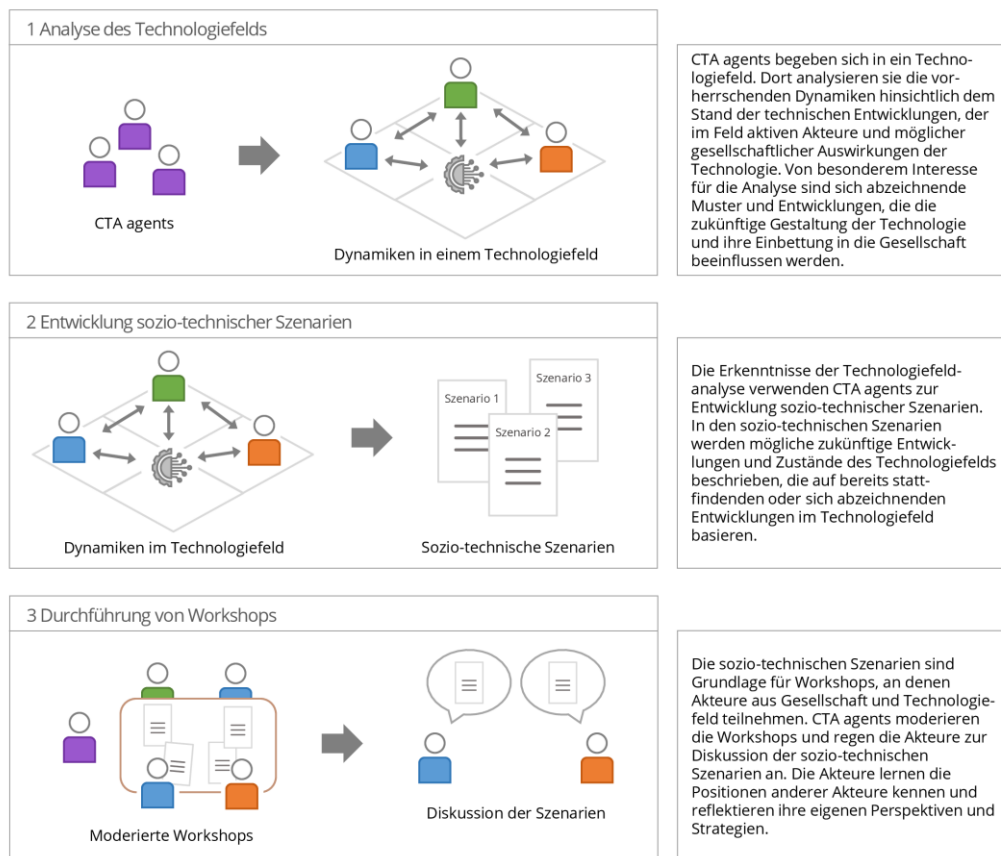
Telefon: +49 721 608-23839  
Email: [pascal.vetter@kit.edu](mailto:pascal.vetter@kit.edu)  
Web: <https://www.itas.kit.edu>

## Constructive Technology Assessment

### Kurzbeschreibung

Constructive Technology Assessment (CTA) hat zum Ziel, bereits in frühen Gestaltungsphasen die Reflexion über die mit einer Technologie zusammenhängenden gesellschaftlichen Auswirkungen anzuregen und dadurch Impulse für eine soziale Technikgestaltung zu setzen. CTA schlägt vor, den Kreis der am Gestaltungsprozess beteiligten Akteure um gesellschaftliche Akteure zu erweitern, die wirtschaftliche, gesundheitliche, kulturelle oder anderweitige Auswirkungen der betreffenden Technologie erfahren werden, jedoch nicht direkt am Gestaltungsprozess beteiligt sind. Abhängig von der Technologie können solche gesellschaftlichen Akteure beispielsweise Verbraucher, Bürger, Arbeitnehmer, Unternehmen oder weitere gesellschaftliche Gruppen darstellen. Die Methode von CTA sieht vor (1) das betreffende Technologiefeld zu analysieren, (2) sozio-technische Szenarien zu entwickeln und (3) Workshops durchzuführen. Durchgeführt wird CTA in der Regel von CTA agents, die typischerweise einen Hintergrund in Techniksoziologie, Technikfolgenabschätzung oder einer verwandten Fachrichtung besitzen.

### Methode



## Value Sensitive Design

### Kurzbeschreibung

Value Sensitive Design (VSD) hat zum Ziel, in Projekten zur Gestaltung technischer Systeme menschliche Werte zu identifizieren und systematisch in die Gestaltung miteinzubeziehen. VSD schlägt vor, bei der Gestaltung drei Arten von Untersuchungen durchzuführen:

- (1) konzeptuelle Untersuchungen um für das Projekt zentrale Sachverhalte und Werte zu identifizieren,
- (2) empirische Untersuchungen um Zusammenhänge von System, Nutzungskontext und Stakeholderwerten zu ergründen, und
- (3) technische Untersuchungen um zu überprüfen, inwieweit ein System menschliche Werte unterstützt oder beeinträchtigt.

Eine Stakeholderanalyse (konzeptionelle Untersuchung) hat zum Ziel, direkte und indirekte Stakeholder, nicht-intendierte Folgen sowie potenzielle Nutzen und Risiken eines Systems zu identifizieren. Durchgeführt wird die Methode in der Regel von Personen aus dem Projektteam, das für die Gestaltung eines Systems verantwortlich ist.

### Konzeptuelle Untersuchung – Methode Stakeholderanalyse

#### 1 Erstellung von Listen zu direkten und indirekten Stakeholdern



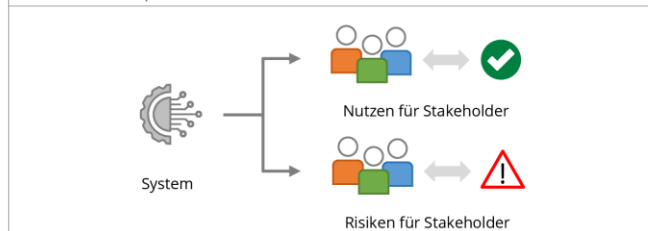
Das Projektteam erstellt zwei Listen, eine Liste zu direkten Stakeholdern und eine Liste zu indirekten Stakeholdern. Direkte Stakeholder sind Personen oder Gruppen, die direkt mit dem System interagieren werden. Indirekte Stakeholder sind Personen oder Gruppen, die nicht direkt mit dem System interagieren werden, die in ihrem Leben jedoch Auswirkungen erfahren werden, die durch die direkten Stakeholder und ihren Interaktionen mit dem System hervorgerufen werden.

#### 2 Identifikation möglicher nicht-intendierter Nutzungskontexte und Folgen



Das Projektteam identifiziert nicht-intendierte Nutzungskontexte und Folgen des Systems. Dies sind Nutzungskontexte, in denen die Nutzung des Systems vorstellbar ist, durch die sowohl positive wie auch negative Effekte hervorgerufen werden können und denen im bisherigen Gestaltungsprozess nur wenig Beachtung geschenkt wurde.

#### 3 Identifikation potenzieller Nutzen und Risiken für Stakeholder



Das Projektteam identifiziert potenzielle Nutzen und Risiken des Systems. Nutzen sind positive Effekte oder Veränderungen für Stakeholder, Risiken sind mögliche Schäden oder Nachteile für Stakeholder, die durch das System entstehen können.

## Value Sensitive Design

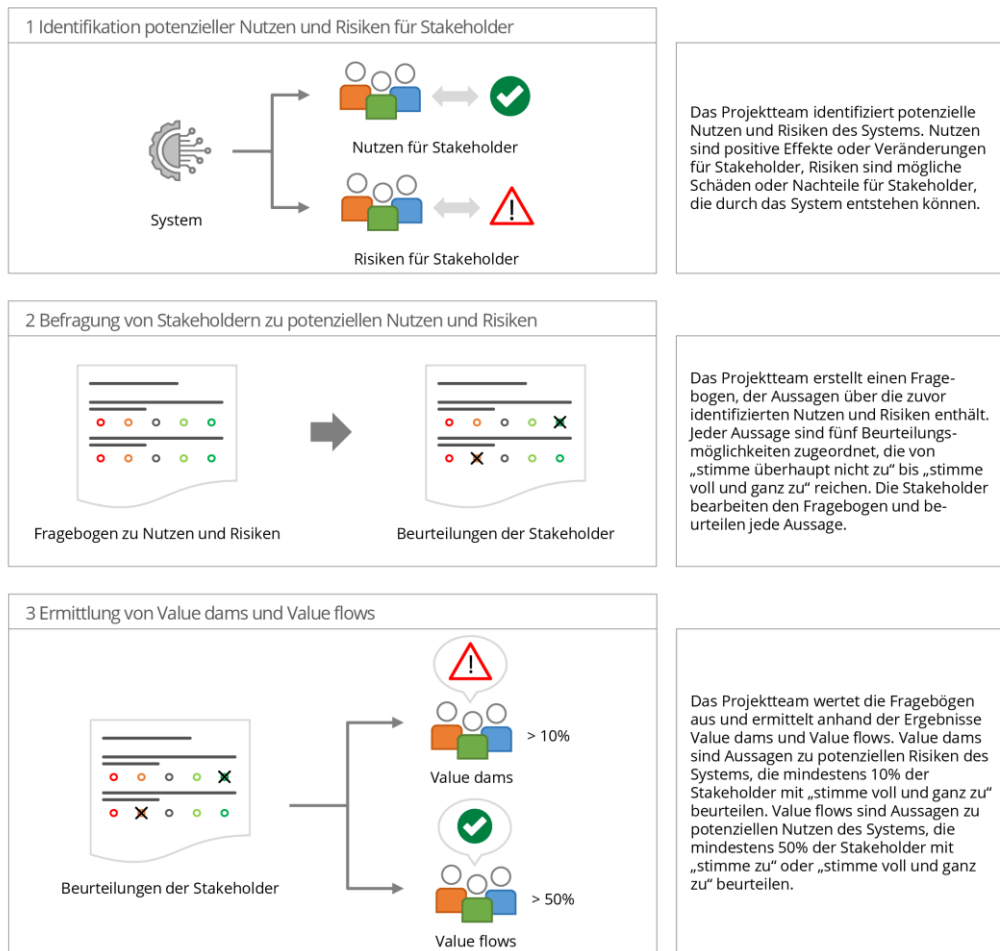
### Kurzbeschreibung

Value Sensitive Design (VSD) hat zum Ziel, in Projekten zur Gestaltung technischer Systeme menschliche Werte zu identifizieren und systematisch in die Gestaltung miteinzubeziehen. VSD schlägt vor, bei der Gestaltung drei Arten von Untersuchungen durchzuführen:

- (1) konzeptuelle Untersuchungen um für das Projekt zentrale Sachverhalte und Werte zu identifizieren,
- (2) empirische Untersuchungen um Zusammenhänge von System, Nutzungskontext und Stakeholderwerten zu ergründen, und
- (3) technische Untersuchungen um zu überprüfen, inwieweit ein System menschliche Werte unterstützt oder beeinträchtigt.

Bei der Methode Value dams and flows (empirische Untersuchung) wird mithilfe eines Fragebogens ermittelt, wie Stakeholder mögliche Nutzen und Risiken des Systems beurteilen. Durchgeführt wird die Methode in der Regel von Personen aus dem Projektteam, das für die Gestaltung eines Systems verantwortlich ist.

### Empirische Untersuchung – Methode Value dams and flows



## Value Sensitive Design

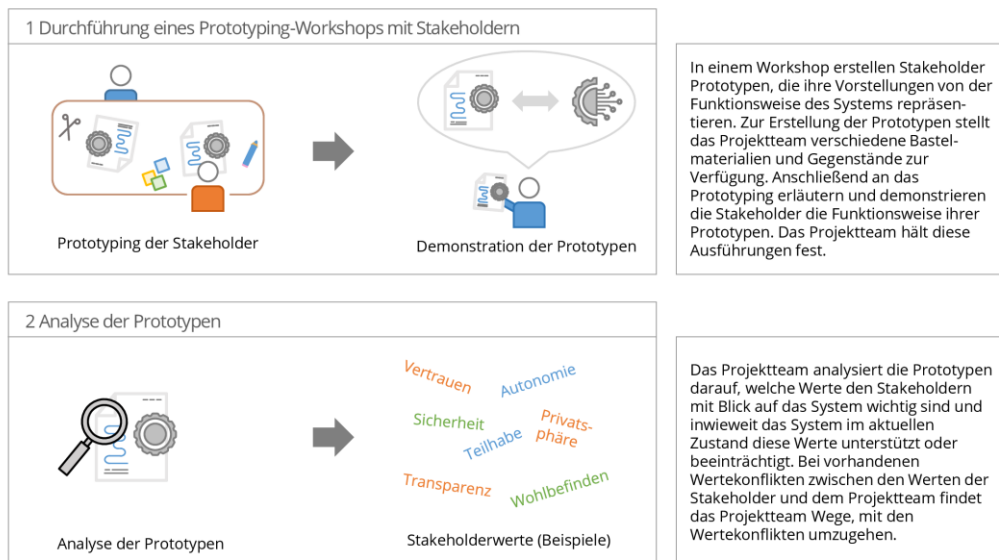
### Kurzbeschreibung

Value Sensitive Design (VSD) hat zum Ziel, in Projekten zur Gestaltung technischer Systeme menschliche Werte zu identifizieren und systematisch in die Gestaltung miteinzubeziehen. VSD schlägt vor, bei der Gestaltung drei Arten von Untersuchungen durchzuführen:

- (1) konzeptuelle Untersuchungen um für das Projekt zentrale Sachverhalte und Werte zu identifizieren,
- (2) empirische Untersuchungen um Zusammenhänge von System, Nutzungskontext und Stakeholderwerten zu ergründen, und
- (3) technische Untersuchungen um zu überprüfen, inwieweit ein System menschliche Werte unterstützt oder beeinträchtigt.

Beim Value-oriented Prototyping (technische Untersuchung) erstellen Stakeholder Prototypen des Systems, die ihre Wertvorstellungen repräsentieren und ihr Verständnis der Funktionsweise des Systems veranschaulichen. Durchgeführt wird die Methode in der Regel von Personen aus dem Projektteam, das für die Gestaltung eines Systems verantwortlich ist.

### Technische Untersuchung – Methode Value-oriented prototype



# Storyboards (englisch)

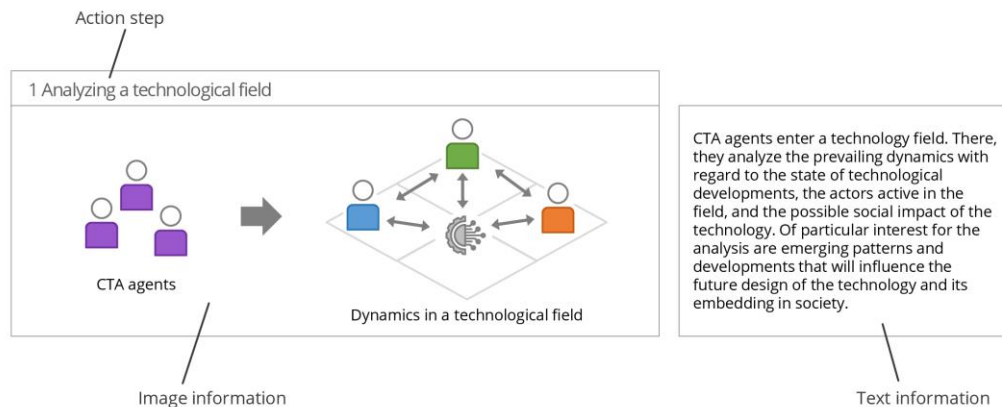


## About this document

This document is part of a scientific investigation and is intended to prepare and support the conduct of expert interviews. The information in this document is primarily addressed to experts and professionals from the field of Artificial Intelligence (AI) who have project experience in the design and development of AI systems.

In this document, design methods of the approaches Constructive Technology Assessment (CTA) and Value Sensitive Design (VSD) are depicted and explained. A combination of image and text information is used to illustrate the methods and their respective action steps (see example below). All information used for this purpose is based on scientific literature.

## Example of information depiction in the document



If you have any questions or comments about this document, please contact:

Karlsruhe Institute of Technology (KIT)  
Institute for Technology Assessment and Systems Analysis (ITAS)

Pascal Vetter, B.Eng., M.Sc.  
Scientific staff

Karlstrasse 11  
76133 Karlsruhe, Germany

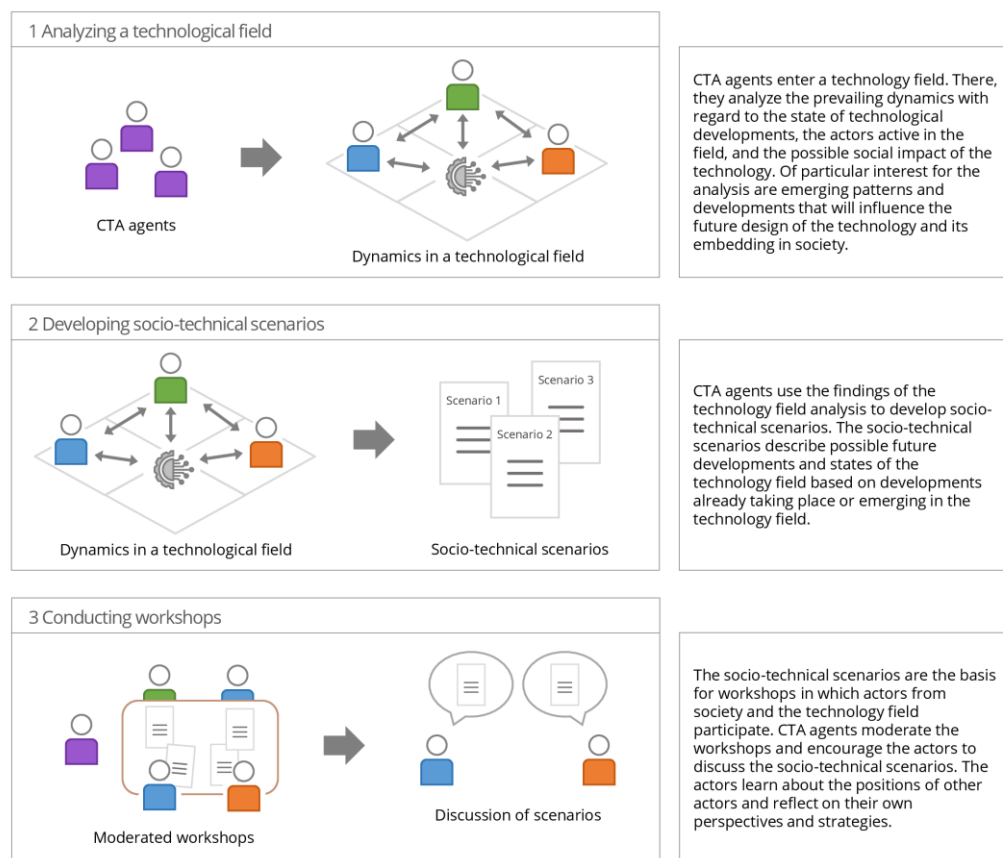
Phone: +49 721 608-23839  
Email: [pascal.vetter@kit.edu](mailto:pascal.vetter@kit.edu)  
Web: <https://www.itas.kit.edu>

## Constructive Technology Assessment

### Short description

Constructive Technology Assessment (CTA) aims to stimulate reflection on the societal impacts associated with a technology at early design stages, thereby providing impetus for social technology design. CTA proposes to expand the group of actors involved in the design process to include societal actors who will experience for example economical, health-related, cultural, or other impacts of the technology in question but are not directly involved in the design process. Depending on the technology, such societal actors may include, for example, consumers, citizens, workers, businesses, or other societal groups. The method of CTA is to (1) analyze the technology field in question, (2) develop socio-technical scenarios, and (3) conduct workshops. CTA is usually conducted by CTA agents, who typically have a background in sociology of technology, technology assessment, or a related field.

### Method



## Value Sensitive Design

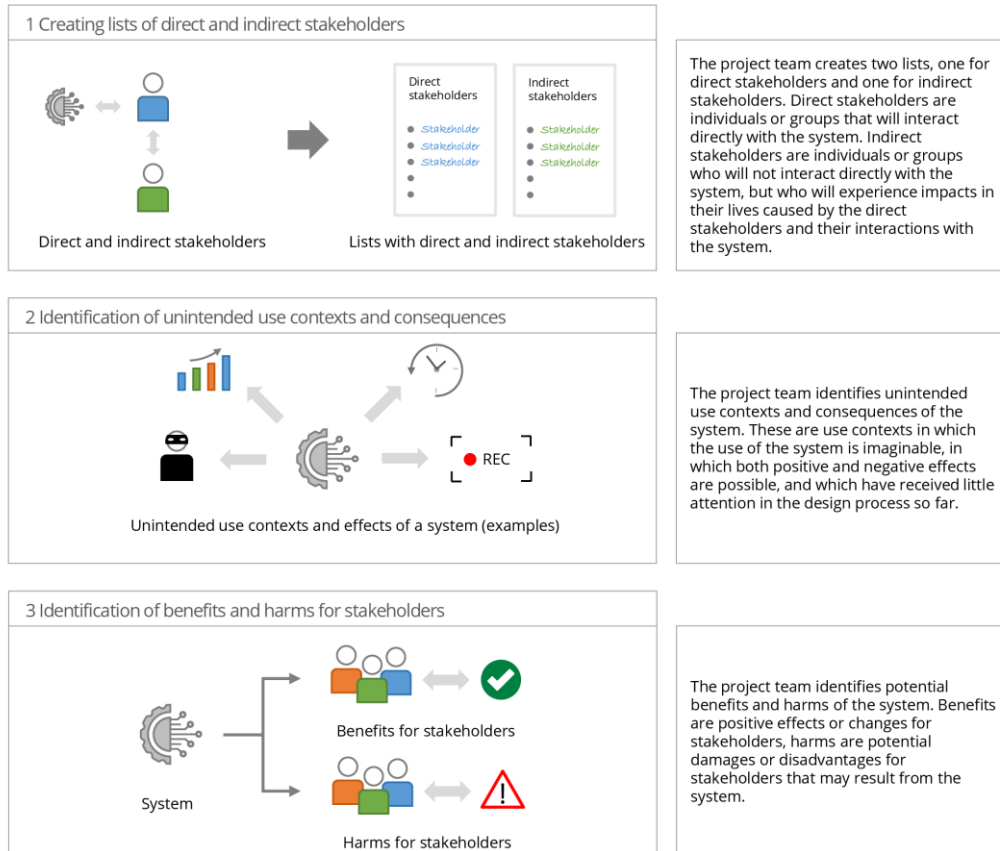
### Short description

Value Sensitive Design (VSD) aims to identify and systematically incorporate human values in projects where technical systems are designed. VSD suggests conducting three types of investigations during the design:

- (1) conceptual investigations to identify issues and values central to the project,
- (2) empirical investigations to determine coherences among system, context of use, and stakeholder values, and
- (3) technical investigations to examine the extent to which a system supports or hinders human values.

A stakeholder analysis (conceptual investigation) aims to identify direct and indirect stakeholders, unintended consequences, and potential benefits and harms of a system. The method is usually carried out by people from the project team responsible for designing a system.

### Conceptual investigation – Method stakeholder analysis





## Value Sensitive Design

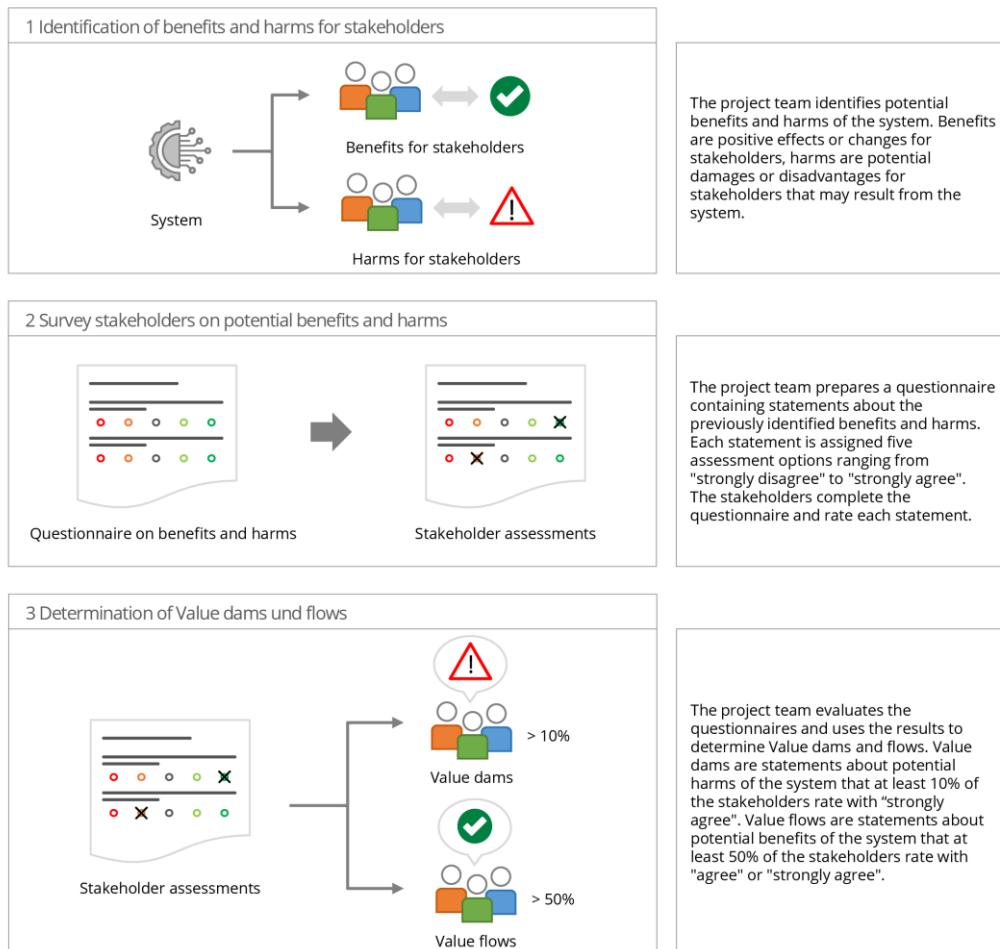
### Short description

Value Sensitive Design (VSD) aims to identify and systematically incorporate human values in projects where technical systems are designed. VSD suggests conducting three types of investigations during the design:

- (1) conceptual investigations to identify issues and values central to the project,
- (2) empirical investigations to determine coherences among system, context of use, and stakeholder values, and
- (3) technical investigations to examine the extent to which a system supports or hinders human values.

The Value dams and flows method (empirical investigation) uses a questionnaire to determine how stakeholders assess the possible benefits and harms of a system. The method is usually carried out by people from the project team responsible for designing a system.

### Empirical investigation – Method Value dams and flows



## Value Sensitive Design

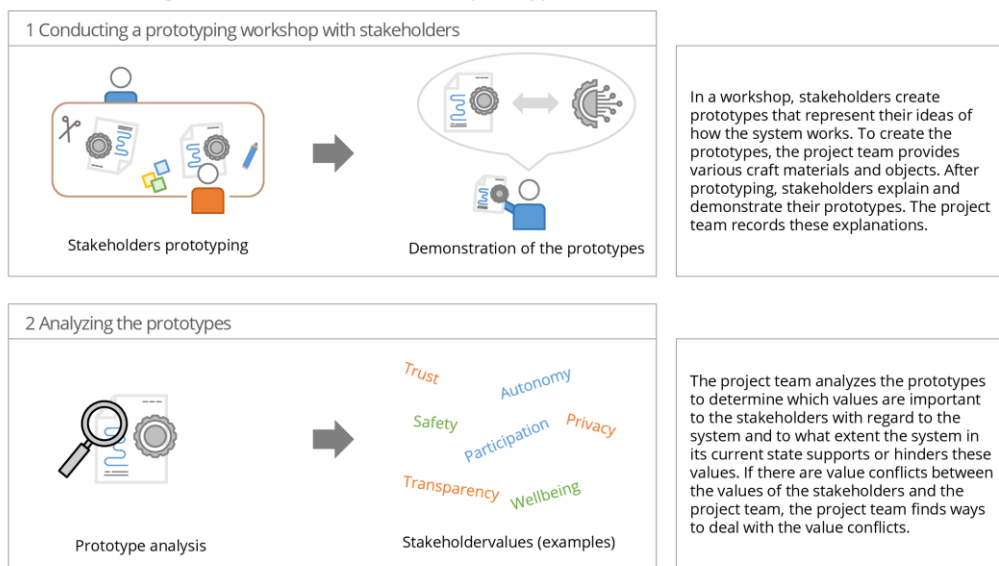
### Short description

Value Sensitive Design (VSD) aims to identify and systematically incorporate human values in projects where technical systems are designed. VSD suggests conducting three types of investigations during the design:

- (1) conceptual investigations to identify issues and values central to the project,
- (2) empirical investigations to determine coherences among system, context of use, and stakeholder values, and
- (3) technical investigations to examine the extent to which a system supports or hinders human values.

The method value-oriented prototyping (technical investigation) involves stakeholders in creating prototypes of the system that represent their values and illustrate their understanding of how the system works. The method is usually carried out by people from the project team responsible for designing a system.

### Technical Investigation – Method Value-oriented prototype



# Interviewleitfaden (deutsch)

## Interviewleitfaden – Experteninterview Evaluation CTA & VSD

### Rahmendaten zum Interview

Gesprächspartner:	
Position und Unternehmen:	
Datum und Ort:	
Uhrzeit:	

### Vorbereitung und Einleitung

- Sinngemäß: „Vielen Dank, dass Sie sich heute Zeit nehmen um mit mir über Gestaltungsmethoden für KI-Systeme zu sprechen.“
- Hinweis auf Pseudonomisierung des Gesprächs
- Hinweis auf Tonaufzeichnung
- Kurze Erläuterung der Dissertation:  
In meiner Dissertation am KIT/ ITAS untersuche ich die beiden Gestaltungsansätze CTA und VSD und möchte herausfinden, inwieweit ihre Methoden geeignet sind, um KI-Systeme menschenzentriert zu gestalten. Dazu spreche ich mit verschiedenen Experten, die Erfahrung mit der Gestaltung von KI-Systemen besitzen. Sie sind einer der Experten und ich freue mich auf unser Gespräch.

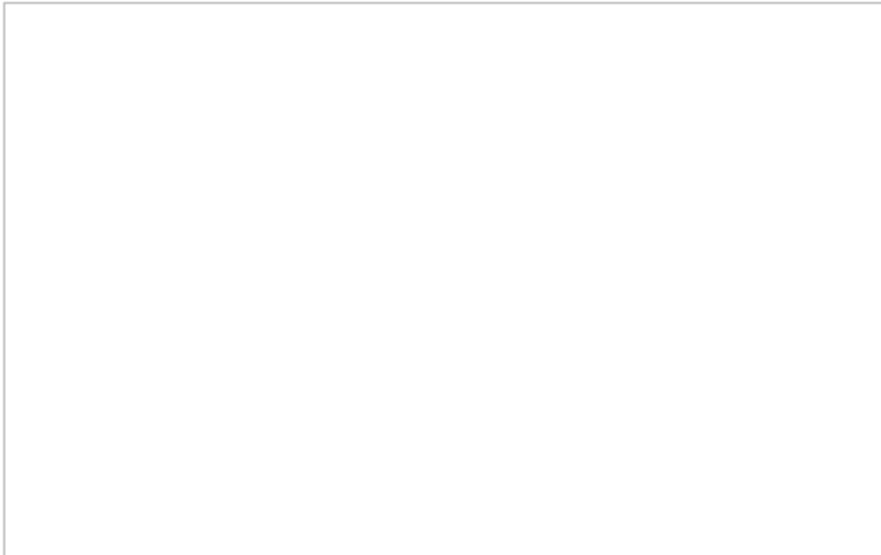
### Erfahrungen aus Projekten zu KI-Gestaltung

Wie waren Sie bislang in die Gestaltung von KI-Systemen involviert? Was waren Ihre Aufgaben?

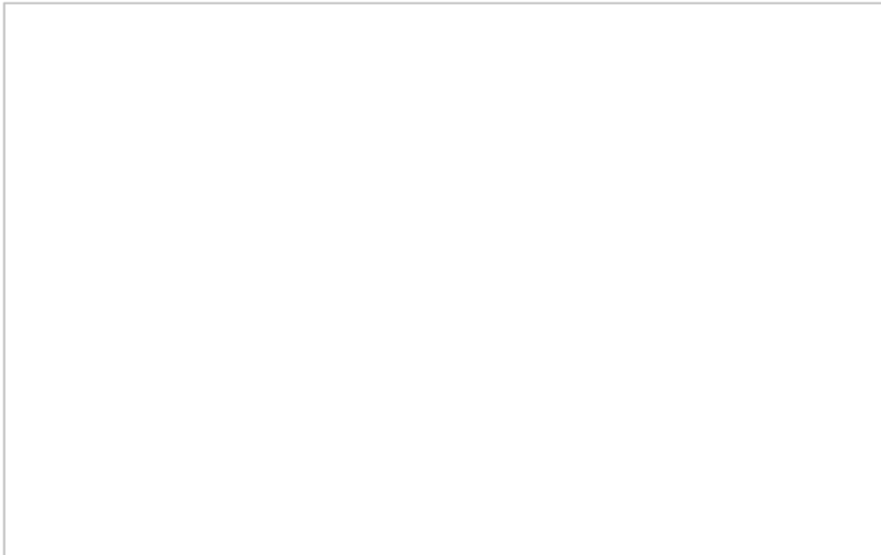
-- Screenshare: Storyboard mit kurzer Erläuterung, (max. 5 Minuten!) --

#### Einschätzung der Methoden von CTA und VSD

Wie beurteilen Sie die Gestaltungsmethoden von CTA und VSD? Was fällt Ihnen auf?

A large, empty rectangular box with a thin black border, intended for the user to provide their evaluation of CTA and VSD methods.

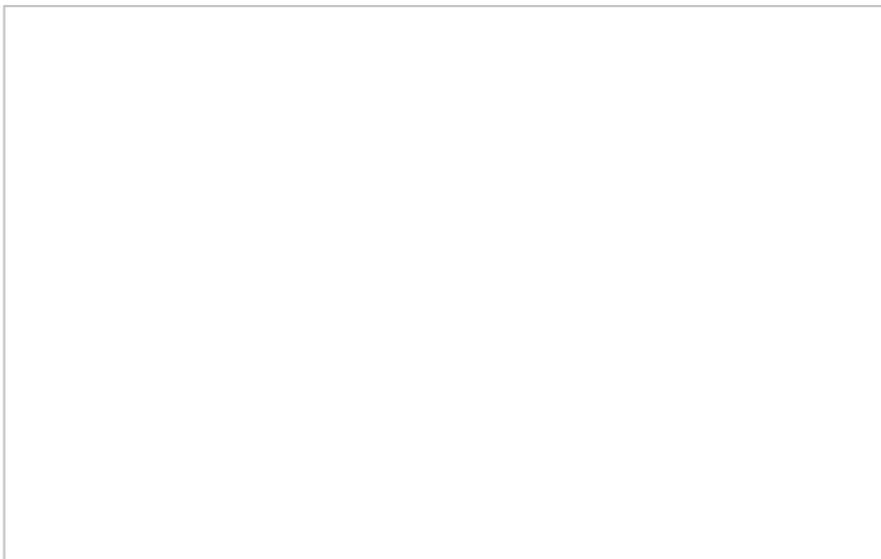
Welche Elemente der Methoden können für die Gestaltung von KI-Systemen verwendet werden?

A large, empty rectangular box with a thin black border, intended for the user to list elements of the methods that can be used for the design of AI systems.

Was können die Methoden für die Gestaltung von KI-Systemen nicht bereitstellen?

A large, empty rectangular box with a thin black border, intended for the user to write their answer to the question above.

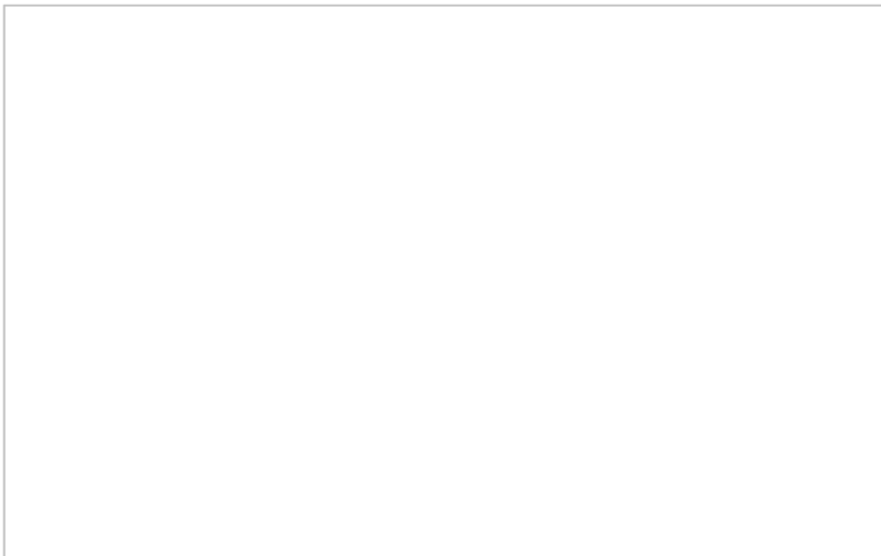
Was begrüßen Sie an den Methoden?

A large, empty rectangular box with a thin black border, intended for the user to write their answer to the question above.

Wo sehen Sie Lücken oder blinde Flecke bei den Methoden? Was wird nicht berücksichtigt?



Stellen Sie sich vor, Sie würden in Ihrem nächsten KI-Projekt mit diesen Methoden arbeiten – Was macht das mit Ihnen?



Möchten Sie noch weitere Gedanken äußern?

### Fragenspeicher

Mit welchen Methoden haben Sie in der Vergangenheit gearbeitet, um KI-Systeme zu gestalten?

Welche Gemeinsamkeiten und Unterschiede bestehen zwischen CTA/ VSD, und den Methoden, mit denen Sie bislang KI-Systeme gestaltet haben?

Was vermissen Sie bei den Methoden, die bislang zur Gestaltung von KI-Systemen eingesetzt werden?

Was müssen Ihrer Einschätzung nach Methoden zur Gestaltung von KI-Systemen leisten können?

# Interviewleitfaden (englisch)

## Interviewleitfaden – Experteninterview Evaluation CTA & VSD

### Rahmendaten zum Interview

Gesprächspartner:	
Position und Unternehmen:	
Datum und Ort:	
Uhrzeit:	

### Vorbereitung und Einleitung

- Sinngemäß: „Thank you for taking the time to talk to me today about design methods for AI systems.“
- Reference to pseudonomization of the conversation
- Reference to audio recording
- Brief explanation of the dissertation:  
In my dissertation at KIT/ ITAS, I am investigating the two design approaches CTA and VSD and investigate to what extent their methods are suitable for designing human-centered AI systems. To this end, I am talking to various experts who have experience with the design of AI systems. You are one of the experts and I look forward to our conversation.

### Erfahrungen aus Projekten zu KI-Gestaltung

How have you been involved in the design of AI systems so far? What were your roles/tasks?



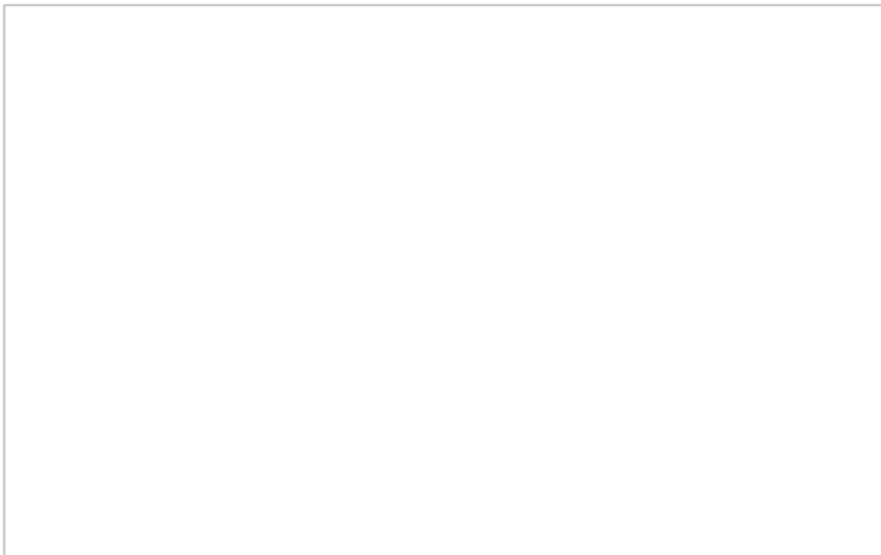
-- Screenshare: Storyboard mit kurzer Erläuterung, (max. 5 Minuten!) --

Einschätzung der Methoden von CTA und VSD

How do you assess the design methods of CTA and VSD? What strikes you?

A large, empty rectangular box with a thin black border, intended for a storyboard or assessment.

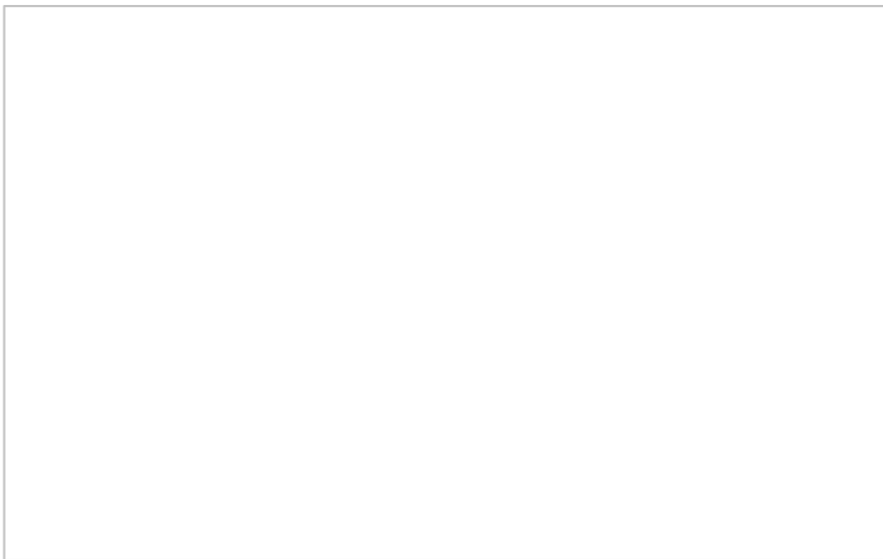
What elements of the methods can be used to design AI systems?

A large, empty rectangular box with a thin black border, intended for design elements.

What can the methods not provide for designing AI systems?

A large, empty rectangular box with a thin black border, intended for the user to write their answer to the question above.

What elements of the methods do you appreciate?

A large, empty rectangular box with a thin black border, intended for the user to write their answer to the question above.

Where do you see gaps or blind spots in the methods? What do they not take into account?

Imagine you had to apply these methods in your current/ next AI-project – How does that make you feel? (Was geht Ihnen dazu durch den Kopf?)

Any other remarks or comments you would like to share?

Seite 4 von 5

### Fragenspeicher

What methods have you used in the past to design AI systems?

What are the similarities and differences between CTA/ VSD, and the methods you have used to design AI systems so far?

What do you miss in the methods that have been used so far to design AI systems?

What do you think methods for designing AI systems must provide?





