Ethics Soc. (2026) 20:25
https://doi.org/10.1007/s11569-026-00499-y

INTERVIEW

Technology Assessment *and* Ethics of Technology? An Insider's Story

Armin Grunwald · Giovanni Frigo

Received: 6 January 2026 / Accepted: 17 January 2026

Abstract This interview with Professor Armin Grunwald, Head of the Institute for Technology Assessment and Systems Analysis (ITAS) at the Karlsruhe Institute of Technology (KIT), offers an in-depth exploration of Technology Assessment (TA) as a conceptual, methodological, and institutional practice. Combining biographical reflection with systematic inquiry, the interview traces the development of Professor Grunwald's intellectual trajectory and examines how his methodological and conceptual perspectives on TA have evolved in dialogue with philosophy, ethics, and related fields. The interview is structured into six thematic parts. It begins with biographical reflections that situate Grunwald's engagement with TA within broader academic and institutional contexts, followed by a discussion of his approach to TA and its theoretical underpinnings. The second part examines the interrelations between TA and other disciplines and approaches, including philosophy and ethics of technology, highlighting both complementarities and enduring tensions. The third part addresses political, social, and economic dimensions of TA, with particular attention to questions of policy advice, democratic governance, and the prospects and challenges of "global" TA. This section foregrounds contextual differences between democratic and non-democratic political systems and engages critically with issues of inclusivity, discrimination, power, and privilege within TA practices. The fourth part connects TA to environmental thinking and environmental ethics, exploring how ecological concerns and sustainability considerations shape and challenge established TA frameworks. The fifth part introduces a set of more provocative questions, including reflections on whether and to what extent artificial intelligence and robotic systems might assume tasks traditionally associated with TA. The interview concludes by turning toward the future of TA, addressing the prospective roles of research institutes such as ITAS and revisiting biographical themes in light of ongoing technological development and innovation. Overall, the interview contributes to a nuanced understanding of TA as a reflective, interdisciplinary, and societally embedded field. The reflections and insights of a leading intellectual and experienced insider are intended to stimulate further dialogue between Technology Assessment and neighboring fields of inquiry and practice.

A. Grunwald · G. Frigo (✉)
Institute for Technology Assessment and Systems Analysis, Karlsruhe Institute of Technology, Karlsruhe, Germany
e-mail: giovanni.frigo@kit.edu

A. Grunwald
e-mail: armin.grunwald@kit.edu

Published online: 06 August 2026

Introduction

As is often the case with engaging discussions, the present interview has evolved over several months into a broader and more nuanced exchange than initially anticipated. Its original aim was twofold: first, to elucidate the philosophical groundings of Technology Assessment (TA), and second, to clarify the conceptual and practical relations between the fields of Philosophy and Ethics of Technology (or Technology Ethics) and that of TA. Against this backdrop, interviewing Professor Armin Grunwald—head of Germany's leading institute for TA[1] as well as of the Office of Technology Assessment at the German Bundestag (TAB)[2] — appeared as a particularly fitting approach. For me, as the interviewer and scholar in Environmental, Technology, and Energy Ethics, posing direct and candid questions to Professor Grunwald has been both an honor and an intellectual delight, helping me appreciate both the historical and cultural developments as well as the conceptual roots of German TA [1] and beyond.

This inquiry is situated within a specific socio-institutional and spatio-temporal context. The broader German university landscape comprises more than 420 higher education institutions, of which approximately 120 are full universities (Universitäten). Within this diverse and rich landscape, around 20 of them are traditional Technische Universitäten (Technical Universities, TUs), institutions whose core mission centers on science, engineering, and technology. Out of the ~20 TUs, 9 are part of the TU9 alliance, a group of the most research-intensive technical universities in Germany. One of these is the Karlsruhe Institute of Technology (KIT), of which both the interviewer and the interviewee are members. Within KIT exists the Institute for Technology Assessment and Systems Analysis (ITAS), which recently celebrated its 30th anniversary.

A noteworthy development at KIT over the past few years has been the attempt to institutionalize TA as a distinct field of study. As a result of these efforts, TA was initially offered as a "minor" subject of study and, since 2025, as a possible "major" within several bachelor's programs, including the Liberal Arts and Sciences (LAS) degree. Through the interviewer's personal involvement in these curricular developments – particularly teaching on the philosophical foundations of TA and on the interrelations between Ethics of Technology and TA—a range of theoretical questions has emerged that merit deeper reflection and more systematic analysis. To obtain some substantial answers, this interview engages with a leading intellectual in the field of TA who can offer an insider's perspective and nuanced observations.

The interviewer's efforts to better understand these issues connect directly to Armin Grunwald's seminal 1999 article, *"Ethics of Technology or Technology Assessment?"* [2]. In this text, Grunwald argues that Ethics of Technology and TA represent two vital yet historically distinct ways of addressing the impacts of technology. When pursued in isolation, each, he suggests, suffers from characteristic deficits. Ethics of Technology, on the one hand, tends to exhibit an "operational deficit," stemming from insufficient empirical and contextual grounding, a limited capacity to address uncertainty, and weak orientation toward policy engagement. TA, on the other hand, risks developing a "normative deficit" when it operates without explicit ethical reflection. In such cases, TA often presents itself as neutral, descriptive, or merely advisory, thereby avoiding the systematic justification of its normative foundations. Ethical questions may be reduced to matters of impact assessment, with deliberation oriented more toward political feasibility than toward moral justification. These deficits and differences motivate Grunwald's broader argument for a closer integration of Ethics of Technology and TA. For robust guidance in technology policy, normative and empirical insights must be brought together rather than kept apart. Given that, since the publication of that 1999 article, the two fields have arguably grown closer and now share more similarities than differences, the title of this interview deliberately echoes Grunwald's earlier contribution by replacing the conjunction "or" with "and." The arguments developed in that article and more recent similar contributions [3], together with the comprehensive perspectives offered in Grunwald's later volume *Technology Assessment in Practice and Theory*

[1] Institute for Technology Assessment and Systems Analysis (ITAS), German: Institut für Technikfolgenabschätzung und Systemanalyse (ITAS).

[2] Office of Technology Assessment at the German Bundestag (TAB), German: Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag (TAB).

[4], provide the conceptual point of departure for the dialogue that follows.

As Professor Grunwald approaches the end of his professional career, the interview seeks to reflect on yesterday while also sensing tomorrow. By the time of his official retirement in 2027, he will have served as head of ITAS for 27 consecutive years (since October 1999). Symbolically, the interview consists of 27 (hopefully stimulating) questions, each followed by Grunwald's response. As a written interview, it includes scholarly references throughout and is intended to be of value both to curious newcomers and to more experienced readers.

Opening of the Interview

Professor Grunwald, thank you very much for agreeing to this interview. It's 2026 so you have been serving as Head of the Institute of Technology Assessment and Systems Analysis (ITAS) for more than 25 years. I would like to ask you a series of questions concerning your career, and in particular your personal path of "thinking through technology" [5]. The interview is structured into six parts.

We will begin with a couple of biographical reflections, followed by questions on how you developed your own methodological and conceptual perspectives in TA and beyond. The second part focuses on the interrelations between TA and other approaches, disciplines, and fields. Part III addresses a set of "political," "social," and "economic" questions. Here, I also aim to discuss issues of "global" TA, with particular attention to contextual differences, for instance between democratic and non-democratic political systems. In addition, this section explores topics such as inclusivity in TA, discrimination, power, and privilege. Part IV comprises questions that connect TA with environmental thinking and environmental ethics. Part V brings together a series of more diverse questions, including several provocative ones concerning the possibility of artificial intelligence and robots performing tasks associated with TA. Finally, in Part VI, I would like to turn to the future of TA and to the prospective status and role of research institutes such as ITAS. The interview concludes by returning to your biography and to the evolution of your thinking about TA and ethics in connection to technological development and innovation.

Much of the conceptual background for this interview is drawn from your book *Technology Assessment in Practice and Theory* (2019), alongside several of your other publications and different contributions by colleagues working in related areas. Some of the questions are more formal, others more institutional or more personal in nature; all are of genuine interest to me, and I hope they will be equally engaging for our readers.

Shall we begin?

Part I

Thinking about Technology: Developing the Grunwald's Perspective

① *GF: You were originally trained as a physicist and have worked for many years in various research institutions. Since October 1999, you have served as Head of the Institute for Technology Assessment and Systems Analysis (ITAS), the largest research institute for Technology Assessment (TA) in Germany, based in Karlsruhe, Baden-Württemberg. In addition, since 2002 you have headed the Office of Technology Assessment at the German Bundestag (TAB), Germany's parliamentary TA institution, and since 2007 you have held the Chair of Philosophy and Ethics of Technology at the Karlsruhe Institute of Technology (KIT).*

Against this biographical and institutional background, how do you understand the interconnections between your professional trajectory and the development of your thinking about technology?

AG: My way to technology assessment and philosophy of technology was absolutely not straightforward. I became a physicist by education and training just because I was fascinated by theoretical approaches such as Classical Mechanics and Quantum Physics. In my PhD thesis, I kept my distance from technology, even to experimental physics. After the PhD, I had to earn money and got a job in the fast-growing software industry. This was, in some respect, a shock to me: from Theoretical Solid-State Physics and the quantum world thrown into information technology and crude applications, e.g. developing software for luggage handling in a new-built airport.

Only then I started university studies in philosophy, as a kind of counterweight against the world of IT. Critical Theory, in particular the *Dialectics of the Enlightenment* by Horkheimer and Adorno, became a new frame of reference to me. And then, something happened: in 1990, the Office of Technology Assessment at the German *Bundestag* was founded. This was made an issue even in our daily (and local) newspaper. Only according to this event, at the age of 30, I learned about the notion of TA. I was immediately interested: based on my patchwork experience in natural science, software industry, and philosophy, so I thought, I could perhaps contribute to that field. I started looking for job opportunities which was, however, difficult because the Internet and all the digital search facilities were not available at that time. I found my first job in TA at the German Aerospace Center. Then, TA became the mission of my professional life.

But you're asking about the influence of this biography on my thinking about technology and TA. Having been a physicist was great – but in my self-assessment this time did not heavily influence my view on technology and innovation. The reason is simple: at that time, I simply was interested in understanding some aspects of nature and finding mathematical formulas to express this understanding. No thought about innovation, no thought about shaping the world by new technology. It was a completely different type of reasoning. However, physics is, as I guess, still part of my thinking in terms of epistemology in the sense that that kind of critical and method-oriented thinking requires a strict obligation of maximal transparency and traceability of reasoning and the process that leads to certain conclusions – which is in close connection to thinking approaches/principles developed in modern philosophy. I think that the high significance of technology and innovation entered my mind during my occupation in the software industry. I was witnessing the dynamic evolution of that field while simultaneously contributing to it. Studying Critical Theory while doing the job in software engineering – having in mind both these worlds simultaneously – might have been a strong influencing factor leading me to, for example, reflections on the ambivalence and sometimes even dialectics of technology and innovation.

② *GF: Your book Technology Assessment in Practice and Theory* [4] *is notable for explicitly placing "practice" before "theory" in its title. You further emphasize this prioritization by arguing that "taking seriously the primacy of practice over theory in TA implies that theorizing technology assessment has to contribute to improving its practice, either directly or indirectly" (p. 6).*

Does this primacy of practice imply that the practice of TA is itself always already a form of theorizing, even when it is not explicitly articulated as such?

How do processes of theorizing relate to the identity of TA as a field—both in terms of its self-understanding and its institutional and epistemic boundaries?

AG: Indeed, TA practices to meet societal challenges come first, historically as well as systematically. TA has been developed and is still developing to respond to demands expressed by policymakers or society, i.e. from outside TA and outside science. The demands usually are rooted in orientation needs facing rapid scientific and technological advances, in uncertainties and controversies about the future with new technologies in various arenas such as the labor market, the health system or mobility. The issues of unintended side-effects, of the ambivalence of technology and the need for technology development in accordance with the *Leitbild* of sustainable development illustrate these challenges quite well. People and policymakers are confronted with several questions and seek knowledge and normative orientation for opinion-forming and decision-making. Various TA *practices* have been developed to serve these orientational needs at various places and in different settings. As an example, TA emerged from practical needs and political diagnoses in the US American Congress, not from theoretical insights or scientific models. This situation is not bound to TA but typical of the field of problem-oriented research, sometimes denoted as post-normal science [6] or mode 2-science [7].

However, you're right that some elements of theory will always have been involved in these practices from their very beginning. Bringing TA to parliaments, for example, makes sense only if an implicit model of the governance of technology in society is presupposed, which assigns a specific role in this overall process to parliaments. Or, another example, for developing prospective knowledge about consequences of technology, often scientific models and theories are used for bridging the time difference between today and tomorrow – this is inevitable, because we don't

have any data from the future. However, in the first example, theory is present only implicitly. Probably, the inventors of TA at the US Congress did not have any explicit theory about the role of the Congress for technology development or dealing with its consequences. The second example shows a more explicit role of theory – but not of a theory of TA. Rather, theories are taken from scientific disciplines, e.g. from economics, engineering or innovation research. Theory of TA has to address a different level: reflect on TA practices for identifying commonalities among different practices, make knowledge types in TA transparent, consider the relation between epistemic and social elements in its procedural nature (assessment), open up spaces for learning and systematizing TA's knowledge, and much more. Without an existing TA practice, any theory of TA would lack objects to reflect upon and would remain void. This is the systematic argument why practice comes first in TA.

Historically, emerging TA practices were quickly made subject to theory and scientific reflection. The early critique on the perceived technocratic bias of TA, expressed by Brian Wynne [8], is a famous example until today. In the meantime, many relationships between theorization and consequences for practice have been established, in particular for proposing new approaches to TA and in TA methodology. The fruitful role of the theory of social constructivism for the development of Constructive TA is probably the most prominent example. Another example is the transformation of results of the philosophy of technology into the practical approach of Value Sensitive Design (e.g., [9]). The many theoretical debates and controversies around the most appropriate approaches, concepts, and methods contributed, together with the evaluation of empirical case studies, to the emergence of a scientific TA community.

③ *GF: In the same book, you introduce what you describe as a "trinity of Technology Assessment," comprising (a) TA as policy advice, (b) TA in public dialogue, and (c) TA in the engineering process. Could you elaborate on how these three dimensions relate to one another in practice and in theory? Furthermore, do you regard any one of them as more fundamental or indispensable than the others for defining the core mission of TA?*

AG: Departing from its origin in the US Congress, the general idea of TA spread to other contexts. TA was adopted in various arenas with different actors, different institutional settings, different tasks, expectations, and challenges. The 'trinity of TA' is a proposal to classify these diverse configurations according to the respective opinion-forming or decision-making setting because this determines the specific expectations towards TA as well as orientates the most appropriate TA concepts and methods. However, these three branches are not disjunct. Participatory approaches, e.g., may be used in all of them, methods like scenario-building have their place in each of them, and concepts like CTA may play a role in TA as policy advice and in public dialogue or shaping technology as well. Even if there is rather little overlap between TA as policy advice and TA as part of engineering because the social arenas, the targets of TA, the expectations regarding advice, and many other context conditions are too different, they share common items like TA's cognitive interest of increasing reflexivity and methods to be applied.

The proposal of the 'TA trinity' addresses a pragmatic issue rather than aiming at strictly classifying entities. Even if many TA elements can be used in all of the three branches, they will be applied in a specific, often strongly context-dependent manner in each individual case with different actors such as engineers, members of parliament, and stakeholders. Fulfilling the demands needs adaptation to the specific context requirements rather than operating in a more general and abstract manner. TA practice, therefore, has to be an *art of appropriate contextualization*. The 'trinity' is, in this perspective, an offer to organize exchange around experiences in TA practice fields along these three actor configurations. The EPTA Network (European Parliamentary Technology Assessment) is an example for this exchange in the field of parliamentary, i.e. policy-advising TA.

TA as *policy advice* still is probably its most prominent and visible branch: providing policymakers with foresight information, ethical reflection, and alternative strategies on how to proceed in the respective field. In this arena, TA does not directly address technology development for applications but considers the policy-relevant boundary conditions of technology development and innovation at the level of public policy and democratic decision-making, such as environmental and safety standards, relation between national and international sustainability goals, consequences for the labour market, and competitiveness of the national economy. Policy options delivered by

TA often address regulatory issues, public funding, incentives for innovation, and strategies for the transformation of infrastructures.

Participatory approaches involving citizens and stakeholders with broader perspectives regarding knowledge and values were adopted since the 1980 s as part of scientific policy advice, on the one hand. On the other, participatory TA also has been contributing to *public dialogue* following the model of deliberative democracy. In this respect, TA has been part of a broad movement towards a more deliberative and inclusive democracy since the 1970 s already. In recent decades, participatory TA also entered the field of TA-supported engineering by the involvement of users and stakeholders at certain stages of the design process. Currently, participatory TA is not only just one TA approach among others but has been penetrating all areas of TA.

TA as *direct contribution to technology development* is close to engineering. Obviously, engineers have to consider consequences of technology: the *intended* impacts in economic and innovation respect, but also *unintended* effects, e.g. possible malfunctions and corresponding liability issues. In this consequentialist mode, engineers are predecessors of today's TA which is reflected, e.g., in engineering ethics [10]. In modern society with advanced technology, however, the traditional approach must be conceptually widened and extended because of complex and systemic interrelations between technology and society. This broader perspective on consequences was taken from cybernetics: a prospective system view on technology and its possible consequences should be able to enrich the engineering process. Design and engineering processes should be informed by systems analytical TA in order to allow for better results with respect to e.g. environmental compatibility or possible technology conflicts. In this line of thought, TA sometimes has been regarded even as a specific kind of systems analysis since the 1970 s [11, 12]. Contemporary life cycle thinking in the environmental and sustainability field follows this system approach.

This approach usually needs also a broader, interdisciplinary approach beyond engineering disciplines. Gene editing, artificial intelligence, sustainability transitions and autonomous driving demonstrate that prospective reasoning must go far beyond engineering and, rather, involve other disciplines like sustainability research, ethics, and social science, e.g. for *shaping technology* due to social expectations and ethical values. Approaches such as Constructive TA (CTA) and Value Sensitive Design (VSD), both addressing the level of concrete products, systems and services, follow this approach, often in cooperation with engineering.

It is difficult, perhaps impossible, to assess the share of TA as direct partner of engineering compared to other TA branches because there is no sharp borderline between TA-informed and other engineering activities. Applying a rather wide and open TA understanding including, for example, any LCA-informed activity in engineering and technology development would, probably, result in regarding the branch of TA in engineering as the quantitatively largest part of the TA trinity. This would mean that the very idea of TA became a quite normal part of engineering, in particular with respect to environmental issues.

④ *GF: You characterize Technology Assessment as acting as an "honest broker" (p. 173), something Roger Pielke has also wrote about at length in the US context. How should this role be specified in concrete TA practices? In what sense can—or should—TA professionals and scholars be understood as honest brokers, particularly with regard to neutrality, normative commitments, and the mediation of diverse interests and forms of expertise?*

AG: TA practice fields involve several actors and actor groups in various roles: decision-makers, scientists, lobbyists, experts, promoters, incumbents, stakeholders, citizens, CSO movements, and so forth. Depending on the respective social configuration and the expectations, TA may or has to take different roles in the process. Therefore, role models of TA haven been repeatedly discussed in the TA debates [4]. Few examples of opposite poles of role assignments are: TA as scientific observer and analyst keeping distance from engagement versus TA as an intervening force struggling e.g. for a sustainability transformation [13], TA as neutral moderator in case of diverging interests versus TA as referee by determining itself what is a 'better technology' in the respective context, thereby corresponding to a 'TA knows best' attitude; TA as Cassandra for early warning against possible risks caused by new technology, e.g. following the heuristics of fear proposed by Hans Jonas [14] versus TA as tracker searching for opportunities delivered by new technology [15].

In my view, the cognitive interest of TA – enhancing reflexivity – and its commitment to democracy leads directly to the necessity of thinking in alternatives [16]. This position clearly corresponds to the role of the honest broker in the classification provided by [17], strictly preventing the risk of TA running into an arrogant and expertocratic 'TA knows best' attitude [4]. Because of the same reason, restricting TA to either Cassandra or tracker would be a misleading dichotomy. For the sake of illustration, I like the metaphor of TA and related policy-advising approaches being mapmakers of the future [18], regarded as an open space of opportunities and alternatives (see also Q17 below). Such road maps demonstrate how future consequences of technology could affect future society [19] and could lead to various developments. TA enriches the alternative pathways and roads into the future sketched by prospective information on their properties and implications like costs and benefits, possible conflicts, sustainability effects, dual-use configurations, and so on. The emerging portfolio of alternative pathways with different properties then can be used as orientation by society and decision-makers. It will orientate but not determine the decisions to be made. In this way, the reflexivity of decision-making on where and how to travel into the future will be increased. Policymakers, managers, engineers and the public can use the maps of the future as orientation for deliberating different pathways to the future and to finally determine, where the travel shall go [18]. Explicit or implicit expressions of the role of the honest broker can be found in many examples from TA fields of practice. For example, the Office of Technology Assessment of the German *Bundestag* (TAB) does not deliver categoric recommendations to the parliament like often academies of science do following the attitude 'science knows best' but rather provides the parliament with alternative options in the sense of the roadmap metaphor mentioned.

However, one question concerning the role models remains: the tension between TA as distant scientific observer and TA as an intervening and engaging force. On the one hand, the role model of the honest broker forces TA to remain distant and neutral. On the other, TA is expected to 'make a difference' in real-world issues and to create impact [20]. There seems to be a latent but also persistent inconsistency within TA. Indeed, this is a tricky constellation: How can TA simultaneously be scientific and transformative?

My approach to solve this supposedly aporetic situation is departing with an image of TA as an honest broker and then asking for its possible impact in this role model. The first observation is that already agenda-setting by TA can intervene into real-world issues. Already raising questions, highlighting items and determining themes for TA exercises may create impact. The famous 'blackout' study of the German TAB [21] is an example which had an impact before publishing results. Second, also developing alternative options, characteristic for the honest broker, may be an intervention. Having more options on the table opens up the spectrum of shaping future developments, e.g. for energy supply or adopting AI systems and allows a broader comparatist view. Overcoming the widespread TINA syndrome (there is no alternative) by pointing to alternatives has transformative power because it offers the opportunity to engage for or against specific options. We could say that TA, even when acting as honest broker, has a subversive dimension when it unmasks TINA-like positions as merely interest-guided or biased by demonstrating that alternatives exist. A further transformative effect of providing alternative options is that the previously existing options will look differently in the presence of new ones because new comparisons or combinations become possible. Third, TA does not only present options but also assigns properties to these options and gives arguments pro and contra as result of the assessment process. The assessment includes evaluative steps based on normative criteria, e.g. of sustainable development or responsible innovation. In light of the arguments, the alternatives are not equal but differ, e.g. with respect to sustainability criteria. Therefore, these arguments developed as an honest broker can have transformative power. If the arguments would be convincing to actors, stakeholders, and advisees, they can indeed make a difference and create impact.

However, these considerations are bound to a crucial presupposition: that arguments can play a significant role in modern society, e.g. in opinion-forming and policy-making. Some may regard and criticize this view as over-idealistic. My conviction is, however, that society does not only consist of arenas with actors fighting against each other but that

argument-based deliberation not only shall play a role but again and again factually does play a role. In accordance to what I said about TA's roots in American pragmatism and deliberative democracy, TA is, as I believe, argument-based [22]. This is a normative conviction which, nevertheless, must not ignore that power-dominated arenas, interests and lobbyism play strong roles in the practice fields of TA and in modern technology conflicts.

⑤ *GF: Technology Assessment encompasses a variety of approaches and forms—such as parliamentary, constructive, discursive, and participatory TA. To what extent is the choice among these approaches determined by the goals of a specific assessment, and to what extent is it shaped by contextual factors such as institutional settings, political cultures, or stakeholder configurations? Finally, do you consider any particular approach to be especially recommendable or paradigmatic for TA, and if so, on what grounds?*

AG: TA was neither developed out of the blue not it is an academic invention. Rather, its emergence shall respond to challenges accompanying the scientific and technological advance. Demand for TA usually is expressed if orientation is needed for dealing with the scientific and technological advance. This advance creates a lot of urgently wanted outcomes but also raises questions because of ambivalences, ambiguities, controversies and conflicts involved in exploiting its benefits but even more in dealing with risks, unintended side-effects, and winner/loser-constellations. The demand-orientation, therefore, belongs to the foundations of TA. These demands may differ heavily depending on context. Addressing issues of the digital transformation in cooperation with computer scientists, taking care of critical national infrastructures in serving parliament's needs, contributing to sustainable development in the transformation of an urban quarter, doing prospective systems analyses for supporting the energy transformation, or reflecting on the dissolution of boundaries between human life and technologies in cooperation with ethics and anthropology: these and many other examples illustrate the diversity of settings where TA is expected to deliver orientation. TA has established different inter- and transdisciplinary practices and methodologies for serving various settings covering parliamentary TA, public dialogue and engineering. Hence, the choice of a specific TA concept to be applied and of the methods to be used must match the respective context. This choice must not depend on what I or the respective TA team wants to get out of the assessment but rather what knowledge and orientation is needed in the respective field. Therefore, clarifying the expectations but also limitations of a TA study together with the demanders and addresses is of major importance before starting the exercise but also requires adjustments throughout the course of the project [20].

The strong demand- and customer-orientation of TA implies a comparably strong degree of contextualization [23]. From the standpoint of classical scientific disciplines, this could be regarded as strange. Depending strongly on demander's expectations and context conditions could create the impression that TA does not work in scientific independence but similar to consultants offering their service to various customers in a free market. I concede that there can be a risk in specific cases. However, TA could and can only establish reputation, e.g. in parliaments, Technical Universities or among scientific disciplines, by doing its work based on a sound scientific fundament, in particular with respect to methodology. Social Epistemology can be regarded as an appropriate approach for acknowledging the specific nature of TA as a 'post-normal science' [6] while taking care of its epistemological soundness [4]. Apparently, TA was successful in gaining reputation by addressees and scientists as well. The latter is indicated, e.g., by the acceptance of TA in academies of science, e.g. in Germany's National Academy of Science and Engineering (acatech), and by institutional reviews of TA institutes. Nevertheless, reflecting the standard reached and perhaps even improving it needs continuous effort in epistemological and methodological respects.

The key for reconciling the demand orientation of TA with its scientific independence is the distinction between agenda-setting and agenda-elaborating. While the former in TA is strongly influenced or even determined by the demanders seeking orientation, e.g. for dealing with conflicts on biotechnology, the latter (i.e., TA exercises commissioned

and funded for providing orientation) has to be performed according to scientific quality requirements and in scientific independence. This is realized and codified, for example, in the working principles of the Office of Technology Assessment at the German Bundestag (TAB). While setting its agenda is a political topic and therefore on behalf of the parliament, the subsequent TA studies have to be performed with scientific autonomy. This is even stated in the contract between the Bundestag and the scientific institution operating TAB.

⑥ GF: Turning to the concept of "rational TA":

Do you consider this your principal paradigm for Technology Assessment?

Could you elaborate on why and how all three of Rescher's types of rationality are necessary for conducting any rational decision-making within TA, and how they interrelate in practice?

AG: The term 'rational TA' was coined in the late 1990s which was, by the way, a fruitful time for developing new TA approaches. In that decade the constructive TA was born [24], the 'innovation-oriented TA' became an issue in Germany, John Grin and colleagues proposed an 'interactive TA' [25], and participatory TA developed quickly [26]. The background of the concept of a 'rational TA' was a heavy debate between social science and philosophy about the appropriate understanding of TA between normativity and empiricism. The concept itself was developed in the TA institution with the incredibly long name (European Academy for the Study of the Consequences of the Scientific and Technological Advance Bad Neuenahr-Ahrweiler GmbH) under the directorship of philosopher and ethicist Carl-Friedrich Gethmann. As his deputy, I was strongly involved in developing rational TA. Its core is, first, the observation that empirical knowledge is necessary for providing orientation according to the political and societal demands towards TA – but not sufficient. Rather, providing orientation needs normative analysis, reflection, and deliberation. The second conviction is that—in contrast to a positivistic view—normative claims, e.g. conflicts around the acceptability of some technology, are accessible to argument-based reasoning rather than being left with bargaining and a checks and balances approach only. An implication of rational TA is the explicit inclusion of normative ethical inquiry in contrast to other approaches at that time.

Hence, the 'rational TA' approach developed more than 20 years ago is an important cornerstone in developing my view on TA. The underlying concept of a *pragmatic rationality* [27] was derived from American Pragmatism [28, 29], (Continental) methodical constructivism [30] and discourse theory [31, 32]. Following Rescher, rationality assessments of propositions, actions or decisions can be separated into the (cognitive) rationality of propositions, the (practical) rationality of the suitability of actions to realise aims (the means/end rationality) and the (evaluative) rationality of the aims or purposes themselves. This approach is highly relevant and valuable in TA in spite of the fact that it is mentioned explicitly only rarely. Its relevance can easily be demonstrated by recalling the epistemological determination of the types to be used in TA [4]: systems knowledge, prospective knowledge, normative orientation, hermeneutic knowledge, and instrumental knowledge. Elements of systems knowledge obviously are obliged to be checked in the framework of the cognitive rationality. Similarly, this also holds for hermeneutic and prospective knowledge. The latter involves the specific challenge of assessing propositions and statements addressing future developments with regard to validity [33]. Instrumental knowledge considering possible measures to be taken directly corresponds to the practical rationality (which I would preferably call 'instrumental rationality').

The normative orientation asks for the justifiability of the aims, purposes or intentions as indicated by Rescher. The mission of TA, however, includes considering also possible *unintended* consequences of technology-induced or -enabled developments. For illustration: if an AI system is developed towards justifiable purposes this does not imply that also all of the consequences of dissemination and use of the system will be ethically or legally welcome. The recent study of the Germans Ethics Council on AI (2023), for example, raises several issues of possible unintended consequences of AI in various application fields, which could come in conflict with ethical principles. Therefore, evaluative rationality in TA is more complex than in other fields.

Rationality in this pragmatist sense [33] does not claim to arrive at universally valid statements. Any deliberation needs a 'pre-deliberative agreement' among the participants for constituting the ground of

arguing and assessing. The substance of this agreement may differ from situation to situation, e.g. with respect to shared values or goals in a specific field among the actors and parties involved. Pragmatist rationality, therefore, does not only *allow* contextualisation but *postulates* it, far away from any universalism because it involves the question rationality for whom and under what boundary conditions. Pragmatist rationality is *relational*, closely linked to the high significance of contextualization in TA. Rationality in this sense is also *procedural,* rooted in philosophical discourse theory [31], In TA, this aspect can easily be found in the reconstruction of TA as assessment process with clear and traceable steps [4]. Finally, the *reflective dimension* of pragmatist rationality is mirrored even in my definition of TA as a set of socio-epistemic practices serving the cognitive interest of enhancing reflexivity for shaping the scientific and technological advance and the usage of its outcomes [4].

In spite of the close relationship between the basic ideas of the 'rational TA' approach and my current thinking, I have not used the term 'rational TA' for more than ten years, maybe even 15 years. One reason might have been that using the term 'rational' motivated many questions. It is burdened by history, e.g. by former planning optimism, falls under suspicions of technocracy, and is often misunderstood as 'rationalistic'. A second reason could have been that I increasingly tried to avoid assigning attributes to my own approach possibly defaming others. Assigning the attribute 'rational' to one's own position may render other approaches and their advocates as 'irrational', at least in a quick and dirty perception. Such self-definitions run the risk of being used or perceived only as polemical terms.

Summarizing, yes, the development of the rational TA approach was a major step in the genesis of my own view on TA. Thanks for raising this question, Giovanni – it was absolutely fascinating to revisit my own but more than 20 years old publications. Looking back to these 20 years, I'm happy to see some continuity but also development. Major elements of the 'rational TA' migrated and disseminated in different forms of adaptation into the TA community. Traces can be found in the recent reflections on normativity in TA (e.g. [34]; cf. Q8 below), in the many references to deliberative democracy with its roots in discourse theory, and on the regular inclusion of ethical reasoning in TA processes.

Part II

Thinking about Technology across Different Approaches, Disciplines, and Fields

⑦ *GF: In this section, we start with the relation between TA and both philosophy and ethics to then delve a bit more into the links between TA and other approaches or fields. Carpenter* ([35], *576*) *tried to clarify some philosophical aspects of* TA, writing for example:

> "Before providing a typology of TA it is important to offer the following qualification. 'Technology' possesses both collective and distributive senses. Phrases such as 'Western technology,' 'society's technology,' 'industrial technology,' etc. reflect the former usage. A growing body of literature, highly significant for the philosophy of technology, adopts the collective sense. TA, on the other hand, involves the distributive usage. It is always the assessment of a single technology such as electronic funds transfer or the breeder reactor, or a class of technologies such as energy technologies, banking technologies, etc. Additionally, TA is future oriented. Examinations of disparities between present day technologies and currently accepted societal norms are generally foreclosed. TA aims to provide "anticipatory control" which can reduce the negative impacts of new or modified technologies. This restrictive focus of TA has been criticized. It could, however, be argued that while such a restriction is a defensible methodological strategy, a better term for characterizing the enterprise might be 'technology anticipation'.
> " *Is there a(ny) difference between "Theory of TA" and "Philosophy of TA"? Or what does a Philosophy of TA look like?*

AG: At the stage of its beginning in the US Congress, TA was established as a practice field for responding to the needs of policymakers. There was neither an explicit theory for doing this nor was the emergence of TA guided by philosophy. It developed just as practice serving the parliament. Later, TA started addressing also further actors in politics, research and engineering, and civil society up to society at large. This further TA development and

diversification was strongly influenced by practical demands and context conditions, too.

Theoretical reflection on TA by sociology (e.g. [8]) and philosophy (e.g. [35]) were published at its early stage already. These reflections, however, came from outside. They were provided by scholars observing the emerging TA field while not being part of it. This contemplative distance is welcome in epistemic respect and produced valuable insights. The development of CTA in the spirit of social constructivism also took place outside of TA, specifically within the field of the sociology of technology (e.g. [36]).

The task of a TA *theory from inside* is, in contrast, to establish a generalizing perspective based on existing practices, to extract characteristic properties of TA, to identify commonalities among different practices and to open up spaces for learning, systematizing TA's knowledge, and further development (cf. the response to Q2 with regard to the primacy of TA's practice over theory). Therefore, I regard theory of TA as a generalizing and systematizing theory of an already existing and developing set of practices. It includes philosophy (epistemology, ethics, philosophy of technology, anthropology, political philosophy …) but also contributions from other disciplines (political theory, governance research, sociology, ecology …). Therefore, I do not speak of a philosophy of TA but of a theory.

In my mind, theorizing TA is not an end in itself. Rather, theoretical considerations should, at the end of the day, support TA in achieving its mission in the 'real world'. Taking seriously the primacy of practice over theory in TA implies that theorizing TA has to go beyond contributing to community-building and developing teachable TA. In addition, any theory of TA has to contribute to improving its *practice*, either directly or indirectly. At the end of the day, theorizing TA must come back to its practice in order to close the learning cycle of the exercise. In this line I argued that theoretical effort should enhance TA's capability to meet practical challenges by [4]:

- determining the conceptual core of technology assessment by generalizing from many individual cases in practice for the sake of its further development but also for communication with the world outside TA;
- developing a structured map of TA concepts, approaches and methods with regard to their suitability for meeting specific challenges in specific contexts and under specific circumstances;
- providing a structured overview of the different actor constellations in TA and clarifying the respective roles of the various actors, in particular advisors and advisees;
- clarifying and categorizing the various processes of assessment in their ontological, epistemological, and evaluative dimensions, and uncovering the criteria used including their origins and justifications;
- establishing a common understanding of terms and notions widely used in TA, such as technology, consequences, risk, responsibility, acceptance, uncertainty, innovation, legitimization, and so forth;
- deriving and systematizing quality criteria of 'good' TA procedures, projects, and results;
- systematizing the knowledge and experience gained by technology assessment historically and through many different practices for making it teachable and usable for community-building.

While I have been often talking or writing about 'philosophy *in* TA', e.g. epistemology, ethics, and hermeneutics, I never used the term 'philosophy *of* TA' so far. My spontaneous understanding of this term is that a philosophy of TA should be part of philosophy of science focusing on the specific nature of TA as scientific but impact-oriented research. As I understand it, a philosophy of TA should elaborate on this specificity with respect to TA knowledge (in particular to its prospective dimension), to the normativity involved (ethics and political philosophy), to relations between humans and technology (anthropology and philosophy of technology), to understanding specific TA contexts (hermeneutics), to an ontology of TA (e.g. concerning its objects as 'absent', [4]), and addressing the very nature of 'assessment' in TA. Partially, these philosophical considerations were included in what I called 'theory of TA' above. However, philosophy of TA might exceed inductive theory-building by relating TA to broader philosophical issues.

⑧ *GF: I now have a longer question, in which I am citing extensively from an early text of yours. In your 1999 article "Technology Assessment or Ethics of Technology?"* [2] *you wrote that*

"These two branches are based on fundamentally different assumptions concerning how to give orientation to technology policy: the philosophical ethics, of course, emphasizes the normative implications of decisions on technology and the importance of moral conflicts (Grunwald, [37]), while technology assessment relies mainly on (descriptive) sociological or economic research (Petermann, [38])", And you argued that the "crux of this conflict between ethics and TA consists mainly of two items: operationalization and normativity" and that the "accusation of TA against the ethical approach to technology is primarily that the ethical approach has an operational deficit and does not meet the demand that technology policy requires concrete consultation." In other words, as you write, "the ethical approach does not know enough about the `real world', about technology and the processes of technological development and about technology policy decisions." TA, "though knowing much about the world of technology, cannot give orientation because of normative deficits". For scientific consultation on technology policy to be successful, "it must be operational as well as normative". Ethical reflection must be "integrated into already existing or still to be developed `practices of technological development in society and there be given a pragmatic site to ensure practical relevance". Practically relevant ethics should be advisory, and "the two approaches are more complementary than contradictory". You wrote: "Of course, the scientific deliberation of political institutions or the public must be normative (at least in questions affecting moral conflicts) as well as operational if it is to be successful." Concerning the normative aspect, you argued: "The German parliamentary TA bureau, in accordance with its statutes, has to renounce prescriptive recommendations and evaluations and to restrict itself to the elaboration of alternative options for decision making. In practice, however, it is not possible to strictly differentiate between description and prescription because the choice of a terminology for making distinctions and the choice of central notions with their (partly normative) connotations already lead to prescriptive aspects and presuppositions of descriptions.". So, one cannot really separate them. "Categorizing ethics as a reflective discipline for factual morals can achieve practical relevance for technology development only if technology decisions involve normative conflicts. If there is no conflict on moral questions, ethical reflection is dispensable and TA may be restricted, for example, to cost/benefit analyses.".

A non-normative TA would thus be quite limited. How do you think today about this relationship between TA and ethics of technology?

AG: Indeed, the 1990s were a decade full of intense debates on how normative and evaluative issues should be taken into account in TA. Philosopher, ethicists, promotors of participatory TA, and social scientists took divergent positions (see [2]). At its philosophical core, it was a clash between ethics claiming that normative aspects of TA were accessible to rational argumentation (see Q4 and Q6), and social sciences assuming that evaluative statements were merely subjective ('Positivismusstreit' in German philosophy).

In the meantime, however, things have been changing considerably. Currently, there is wide consensus that TA inevitably has to deal with normative issues, and that TA should do this in an explicit and transparent manner, e.g. in the field of sustainable development, energy justice, genome editing, or regarding the future of human/machine interfaces. TA cannot escape from doing normative evaluations at some stages in its assessment processes. It needs applying evaluation criteria in particular in case of moral conflicts. Such conflicts often go beyond the instrumental dimension of technical means and instruments for ends and purposes. The fields of biotechnology, medical technologies and life sciences provide us with many examples of moral conflict, e.g. in addressing the beginning and end of human life, the value of biodiversity, or intervening into the germline of organisms and even humans. Sustainable development is another field of conflicts, e.g. concerning the distribution of burdens due to climate protection among nation states, parts of society, and/or generations. The notion of conflict here includes moral ambiguities, uncertainties, and indifferences, as well as new challenges for which moral customs have yet to be

established, or where there are doubts as to whether established moral traditions apply.

Consequently, the questions for evaluation criteria and how they can be justified and receive legitimization have to be answered. At this point, the above-mentioned controversy popped up again. While ethics pointed to normative argumentation (e.g. [39]), the empirical social sciences offered value research for identifying well-accepted evaluation criteria. The idea behind was achieving socially better compatible technology by synchronizing technology development with widely shared public values and attitudes. Ethics, however, warned against a naturalistic fallacy [40] in concluding from an empirical Is to the normative Ought and identified further critical counter-arguments.

Participatory TA argued differently. Following the approach of deliberative democracy based on discourse ethics it was postulated that legitimate evaluation criteria only could be the result of mandated and/or representative participatory processes, in contrast to involving ethics experts for providing professional ethical orientation. The ethical quality of the deliberative and participatory processes in the sense of the discourse ethics should be ensured by applying procedural principles derived from discourse ethics such as fairness, equity, and equal access to information [41].

In contrast to the former conceptual contest, complementary cooperation has been evolving in the meantime. TA today involves expertise from ethics and social science as well while also being participatory. Participative procedures can be informed by research into values and by normative ethics as well. Assessment processes can be understood as operable implementations of discourse ethics [42]. Participation on moral issues cannot be replaced by roundtables of philosophical ethicists only. TA needs inclusion as real deliberation (*wirkliches Gespräch,* cf. [43]). Following the conceptualization of politics, the public, and democracy proposed by American philosopher John Dewey "… democracy is the regulation of the public interest arising from indirect consequences and related conflicting interests; it is combined with the idea that everyone should be involved and, in principle, regarded as a person capable of co-deciding about a regulation of such indirect consequences" [44], I consider ethics as argument-based advice to democratic reasoning, public deliberation, and decision-making. Its target is to empower the respective actors to find better-reflected positions and solutions by themselves. Today's 'peaceful cooperation' of normative and applied ethics, empirical social sciences, and participation needs some common understanding of the assessment process. In my view this comprises:

- philosophical ethics cannot claim to deliver *categorical* conclusions for the morally right actions but is, more modestly, *conditionally-normative* in fields of practical application [45];
- vice versa, the normativity needed in the assessment process cannot solely be taken from an empirical investigation of attitudes and perceptions, as social sciences concede;
- participation and inclusion are needed in order to co-construct legitimate normative orientations;
- transparency in building normative arguments is an inherent commitment of TA.

The creation of normativity in assessment processes rightfully claiming legitimization works pragmatically in many fields and cases. However, there are still a lot of open questions, for example concerning the role of legitimate normativity in making the precautionary principle work [46, 47], the relation between the normativity created in participatory assessments versus the normativity of formally legitimated assessment processes, and legitimacy-creating procedures in case of unresolvable conflicts.

⑨ *GF: Taking a broader perspective, including Science and Technology Studies and Responsible Research and Innovation (RRI), in addition to the Ethics of Technology and TA we have already discussed: how do these approaches, disciplines, or fields differ from one another in terms of their aims, methods, and conceptual frameworks?*

AG: STS is an interdisciplinary research field dominated by sociology but also including historian and philosophical analysis at the borderlines between science and technology, on the one hand, and humans and society, on the other. In its self-understanding, STS takes the epistemological position of a scientific observer and analyst of science and technology in society. It does not aim at directly intervening into real-world issues. As a more distant

observer of the field in the traditional role of social science trying to avoid any intervention, it aims at *understanding* past and ongoing developments by providing explanatory knowledge. TA, however, goes beyond distant analysis. It aims at creating impact and making a difference for real-word issues following a strong demand-orientation and providing 'knowledge for action'.

This sharp difference with respect to self-understanding, objectives, and epistemological roles can be traced back to the roots of STS and TA. While STS was born in the 1970s in the world of academia, the first TA institution, the US American OTA, was established far away from universities: in the American Congress. While STS, therefore, has been developing in academic freedom, TA emerged close to societal actors and depending on respective funds, first by political institutions like parliaments. TA's neighborhood to politics and decision-making, with other words, its institutional position as 'part of the system' while STS claimed to be outside of the political and technological system to be able to criticize it, again and again motivated criticism of STS scholars against TA. Brian Wynne's accuse of TA being technocratic [8] is an early example. In the meantime, theoretical TA debates emphasized the issue of independence. Theorizing TA supports reflections of its own setting in order to avoid misuse and bias as well as to open up spaces for a critical distance. Thinking in alternatives, which is currently frequently used in TA, goes even further: by opening up alternatives, TA could even be subversive in contrast to technocratic approaches of, e.g. 'science knows best' or optimization ideologies of engineers and politicians.

Hence, the relation between STS and TA was and still is characterized by a fruitful tension which motivated learning processes. Today, the former sharp separation of both communities has partially disappeared. The invention of CTA based on insights of STS, sociology and empirical research on the genesis of technology enriched the landscape of TA approaches considerably. Many TA scholars (including me) regularly visit STS conferences, give presentations and receive feedback on overlapping fields of interest, e.g. around the 'sociotechnical imaginaries' proposed by Sheila Jasanoff or on sociology of knowledge and ignorance by Stefan Böschen and others. Vice versa, we often can welcome STS colleagues at TA conferences.

The relation between TA and RRI is completely different. The notion of responsible (research and) innovation was created in US-American and then European research policy for designing a new approach to technology governance. It shall better reconcile science and technology, on the one hand, and social demands and perceptions, on the other. Its interventionist approach is similar to TA. Indeed, TA is one of the main roots of RRI [48] with some new accentuations and shifts of emphasis. In a sense, RRI is a more recent member of the big TA family. RRI intends to involve ethical and social issues more directly in the innovation process by integrative approaches to development and innovation, reflected in the notion of responsibility, to bridge the gap between innovation practice, engineering ethics, TA, governance research and STS, to add responsibility reflections and to support 'constructive paths' of the co-evolution of technology and society. In this respect, RRI can be regarded as a daughter of constructive TA plus explicit ethical reflection [49] into shaping technology and innovation.

Four conceptual dimensions have been proposed for RRI [50]. A more responsible research and innovation shall be diverse and inclusive, anticipative and reflective, open and transparent as well as responsive and adaptive to change. Comparing these dimensions with the three conceptual dimensions of TA – in my version: anticipation, inclusion, and complexity management for enhancing reflexivity (Q6 above) – demonstrates the close neighborhood.

Interestingly, the notion of responsibility was only rarely mentioned in the TA debates. This is surprising because implicitly TA has been motivated by responsibility issues since its beginning (see following Q11). The approach, putting effort into harvesting the positive effects while containing, compensating or avoiding negative aspects is an obvious expression of dealing responsibly with the technological advance. RRI focusing directly on innovation [48] rather than on technology made TA's implicit commitment to responsibility explicit. RRI presupposes that research and innovation could lack responsibility and could even be irresponsible [51] which directly relates to the diagnosis of the *ambivalence of technology and of innovation as well*: both can show unintended effects. Responsibility means taking a look at both sides of the coin

before making judgments and decisions – and this reflexive approach relates RRI closely with TA.

During the 1980 s and 1990 s, the neoliberal wave (Reagonomics and Thatcherism) changed power relations in favor of the economy. The (at least perceived) decreased influence of democratic institutions in the neoliberal era created a 'legitimacy vacuum' [52] with respect to how to determine common interests and the common good. As a hypothesis, I assume that RRI is expected to act as a gap-filler in this respect.[3] In daily practice, however, we see a huge overlap between TA and RRI with respect to conceptual issues and methods while focusing on different parts of the innovation system. TA, if I may say with my TA-bias, is much broader because keeping the political level as an important actor in technology governance and, therefore, as an important addressee of TA.

⑩ *GF: Philosopher of technology Carl Mitcham suggest that scholars who are engaged in philosophy of technology, TA, STS should pay more attention to "engineering practices" (i.e., what engineers do") together with their already established interests in technological devices and systems. Given that you also mention that "TA in the engineering process" is one of the three elements of the "trinity of TA", what do you think about the relationship between TA and engineering (practices)?*

AG: Many philosophers and engineers like the idea that informing engineering processes and engineers in person by ethically proved values and legitimized goals for technology development will help turning the technological advance to the better. This was already the case with the – among German engineers famous – guideline on technology assessment approved by the German Association of Engineers (VDI) in 1991 [53]. In this guideline, values for guiding engineering were determined beyond the dimension of techno-economic optimization like environmental compatibility and supporting the common good. More recent approaches are Design for Value, Value-Sensitive Design, Ethics by Design, and Privacy by Design. The underlying conviction is that, if technology would be designed according to well-reflected and ethically sound values then the outcomes of engineering will show the desired properties and avoid, to a large extent, unsound and unintended effects.

I feel that this assumption is not quite wrong. Of course, it is much better to develop technologies with respect to carefully reflected values and with good intentions than developing technology simply for market purposes, or even without reflecting anything outside of technical feasibility. Therefore, I agree that engineers should have some ethical literacy, and that they should fully be aware that their work is not value-neutral.

My only skeptical point is that the assumption includes some relicts of determinism. The premise is that if good values would be implemented in the respective technology, then unintended side-effects could more or less be avoided. It is exactly this junction which I'd like to draw into doubt because I do not think that a proper design of new technology can determine its later use and the consequences of that use. Unintended consequences may emerge from unexpected ways in which humans use those technologies, independently from the design ideas. Human creativity and flexibility may undermine any deterministic expectation about being able to fix consequences of technology already at the stage of design [54].

Having said this, my conclusion is: cooperation between engineers, ethics, and TA is necessary for many reasons – insofar Carl Mitcham is right – but will not be sufficient for realizing a future without severe, negative consequences of technology. The openness of the future of human action cannot fully be anticipated by responsible engineers but, rather, needs approaches of accompanying the R&D process as well as of monitoring innovations (products, services, systems) and their empirically observable consequences.

⑪ *GF: In the 2019 book you did not expand on TA and RRI (due to space). In 2014 you wrote a chapter "Technology Assessment for Responsible Innovation" where you worked on the relationship between RI and the notion of Responsibility. I was wondering, is Responsibility, in the end, the most important philosophical notion or concept in TA? Should it be considered a principle? If not, which concept are the crucial ones? For example, what is your take on the Precautionary Principle?*

[3] Fulfilling this expectation would be, in my eyes, deeply ambivalent: democratic reasoning and legitimate decision-making can be complemented but not replaced by decentralized reflections on values and responsibility.

AG: This question meets an interesting issue in TA and its history. As far as I know about the emergence and development of TA, there has never been a TA without any background of responsibility. However, differently from RRI (see Q9), the close relation with responsibility has (almost) never been made an *explicit* issue in the TA debates. Probably, responsibility was more a self-evident background issue.

Anyway, the matter of fact is quite clear: the huge impact of today's technology and the spreading of many of its consequences across space and time demonstrate that introducing new technology by following a 'trial and error' approach or a 'wait and see' strategy can no longer be considered responsible. Rather, prospective and, in many cases, precautionary approaches with enhanced reflexivity are needed. The American philosopher with German roots, Hans Jonas, made this a well-perceived issue in the volume 'Imperative of Responsibility' [14]. Resulting questions fuelling TA, therefore, are: How can humankind benefit from the technological advance in a responsible and sustainable manner? How can a society, which places its hopes and trust in innovation and progress, and must continue to do so in the future, protect itself from undesirable and perhaps threatening side effects? How can society preventatively act to cope with possible future adverse effects? How can society increase its resilience? And finally, perhaps a bit too big: how can humankind arrive at a 'good Anthropocene'?

In order to enable managing the scarcity of resources, environmental problems, and the emergence of unintended side effects of the scientific and technological advance, several science- and humanities-based research activities were proposed and are practiced. Socio-ecological research, transition management, sustainability research, and, more recently, RRI are neighbours in this respect [55]. Their overall mission is to extend foresight and reflection capabilities in the field of scientific and technological advance for making best use of its outcomes. Recently, the *Leitbild* of Sustainable Development with its dimension of responsibility over time and across the globe was even determined part of the normative fundament of TA in addition to human rights and democracy [56]. Hence, TA is inherently intertwined with responsibility issues which was made a conceptual issue in RRI decades later only.

In many fields, this relates back to the analysis given by Hans Jonas in a specific way. A major insight of Jonas [14] was that major ethical challenges to modern society do not arise from dys- or malfunctioning technology but rather from technologies smoothly doing their service as intended. Their smooth functioning is, and this is a major motivation of TA, often accompanied by unintended effects. They often appear only after decades, such as the health problems caused by asbestos inhalation [57] or the ozone hole after decades of releasing CCFCs into the atmosphere. Because they are only slowly emerging, such as the loss of biodiversity, they are often difficult to recognize. Often, they lead to heavy societal debates with respect to adequate responses between dramatization and playing down their relevance. Our situation today reads like a confirmation of the diagnosis of Jonas: climate change, to cite just one example, is the result of technology that *works*, e.g., in the form of combustion engines, meat-producing agriculture and the food industry, or coal power plants. Another example is the damage to public debate and threat to democracy in the course of anonymous and fast communication via social media: hate speech, fake news, manipulation, echo chambers which have been occurring because the Internet and social media function well. The significance of this type of slowly emerging unintended effects has been a major motivation for TA with regard to early recognition, early warning, and early evaluation of the data available for weighing chances and risks, or costs and benefits. This approach is directly linked to precautionary reasoning.

TA is therefore closely related to risk research, assessment and management. Several characteristics of the risks of modern technology have to be addressed, such as their global distribution, their far range into the future (e.g. high-level radioactive waste), the potentially extremely high extent and impact of damage (e.g. climate change), complex chains of causes and effects (e.g. leading to the ozone hole), the distribution of responsibility across a huge variety of actors with the risk of a 'thinning' of responsibility, and the issue of irreversible impact (e.g. the eradication of species and loss of biodiversity). In this context of dealing responsibly with risk, TA also relates itself to the diagnoses of a risk society and of reflexive modernization [58, 59]. Beck's central thesis that the scientific and technological

progress dramatically increases the depth of human intervention into the entire world, transforms *hazards* (threats coming from nature or from fate) into *risks* (threats where anthropogenic intervention is part of the cause/effect-chain). This diagnosis implies that human decisions are involved in more and more events and developments as part of the respective cause/effect chains, which, today, can be understood as an expression of the Anthropocene. Together with humanity's increasing capacity for technological intervention into the world also humanity's responsibility increases [14]. TA is among the approaches to open up strategies and pathways to deal constructively with this grand and perhaps frightening diagnosis.

Part III

Thinking about Technology through its Socio-economic and Political Implications

⑫ *GF: Technology Assessment has primarily developed in Western contexts, originating in the USA and subsequently evolving in several European countries. In this sense, it can be seen not only as context-dependent but also potentially partial or culturally specific.*

In your view, what might a non-Western approach to TA look like, and what principles or practices would distinguish it?

How, in your opinion, would such a non-Western form of TA differ from the established Western paradigms in terms of methodology, institutional embedding, or normative assumptions?

AG: Indeed, TA has its origins in Western culture, it was developed and is applied mostly in industrialized countries of the so-called Global North. It is a Western construct exactly because the demand for TA first emerged in industrialized countries, as consequence of their success with technological development and related business models. It is also quite clear that TA must go beyond and reach the global level because technology development and innovation strategies today are to a large share in the hands of global actors in science and engineering and in the economy [60], on the one hand, and because of global challenges such as e.g. climate change, world-wide digitization, resource scarcity, and pollution needs assessments and transformational strategies on the global level, on the other. Going global is, according to my understanding of your question, not simply an issue of disseminating the Western TA model to other parts of the world.

I completely agree with this thesis. However, before going to your both questions, I would like to mention a point possibly causing opposition. Currently I see a tendency to devaluate everything that was developed in the Western world as biased, as parochial and paternalistic, perhaps even as an outcome of white-color egoism in an arrogant attitude considering other positions as of minor relevance and dignity. Without any doubt, this is frequently the case, and traces of colonialist arrogance can still be found in several settings, also neo-colonialism play a role. However, my point is that the fact that something was developed or invented in Europe or the US is not automatically parochial and paternalistic. In a logically forcing manner, it cannot be excluded that there could be a difference between 'developed in Europe' and 'eurocentric' in a paternalistic sense. A more detailed consideration is necessary because Western reasoning and ethics possibly is *not necessarily* eurocentric but may include some more universal ideas.

Regarding TA, I feel that there are indeed some universalizable elements, at least at the more abstract level of its framework [4] with the conceptual dimensions of anticipation, inclusion, and relevance. In my opinion, TA cannot work without any anticipatory elements, without any inclusion, and without reflecting and deciding about relevance. Without these, TA's mission is just neither addressable nor fulfillable.

When leaving this abstract level and coming to more specific issues, however, things may look differently because cultural issues may become relevant. For example, the difference between the Western primacy of the individual versus the Asian emphasis of the collective, or Western egalitarianism versus Chinese emphasis on meritocracy could become relevant issues in assessment. TA's needs for contextualization also concerns cultural issues, historic developments, different governance regimes and also a different understanding of democracy (see Q14). As soon as specific challenges have to be addressed, local traditions and culture become part of the game. They may affect, for example, the understanding of sustainability, the perception of the relation between technology and life, resp. of technology and humans, or ideas of a

good society. TA's conceptual dimension of inclusion not only allows for involving them in the assessment process but, rather, *requires* this involvement. Therefore, determining what TA fitting to cultural context should look like, would be the task of inclusive TA itself. From my perspective, I do neither feel able nor mandated to say anything about how a non-Western TA could look like. This has to be determined by the respective actors, citizens, stakeholders, and policymakers by themselves. Developing a trans-cultural TA only can happen in cooperation with other cultures: TA with them, not for them. Hence, my view is that the TA framework [4] is not eurocentic but includes the perspective of cultural contextualization.

Existing TA, however, could prepare itself for this adaptation by investigating cultural issues such as the perception of technologies and self-images of humans facing new technology as well as issues of the technology governance in other parts of the world, e.g. in countries of the Global South. With respect to cultural differences, the respective languages are an important source of insight. Differences of meaning already occur between languages in Western countries. While the English (American) term TA focuses on technology, the German translation *Technikfolgenabschätzung* addresses its consequences. The French term *Evaluation des Choix Scientifiques et Technologiques* involves evaluative rationality with a strong emphasis on the issue of choice. Different meaning are also assigned to the notions of technology and technique. The Chinese term 技术 (jìshù) goes more in the direction of technique in the sense of a procedure (Chinese) while 'Technik' in German highlights artifacts. I guess, it would be valuable to investigate the meanings and their accentuations in more and in particular non-Western languages to develop a map of different meanings for improving mutual understanding. Hence, intercultural work is required in TA in several respects, ranging from linguistic, cultural, and religious issues to historical experiences, and current perceptions of the global situation and the respective cultural contexts.

⑬ *GF: Building on the idea of non-Western perspectives, how do you view the concept of "global TA" as articulated, for example, in the work of Leonhard Hennen and Julia Hahn* [61]*? In what ways does this framework address the challenges of contextual diversity, inclusivity, and cross-cultural applicability in Technology Assessment?*

AG: TA emerged from national regimes of technology governance. It started at the U.S. Congress and then migrated to European countries. In each country, it was embedded into national traditions of governance and decision-making. Regarding the plurality of European governance regimes, TA developed differently and established developed institutional realizations adapted to the respective systems [62]. In each context, a specific "TA-habitat" emerged [61]. Therefore, each of them first was national and dominated by the respective native language. Already addressing the European level and going for European exchange on TA was a big step in that world. The EPTA network (European Parliamentary Technology Assessment) was an outcome of these first steps. In the last decade, it expanded to the international level by integrating members of e.g. South Korea, Japan, Chile, the United States and Argentina. Hence, we could witness a movement from the previously nation-dominated and separated TA communities toward a more global movement creating a global TA community, in particular in the Global TA Network founded in 2019.

This community is wonderful for exchanging many issues around TA like case studies, experiences, lessons learned, cooperation with scientific disciplines, with civil society, and with policymakers. However, an important element is still missing: there is no global body observing the scientific and technological advance in general, drawing conclusions, and being mandated to act. Apart from these considerations of governance and organizational issues at the global level, any global TA also has in addition to deal with the challenge of doing TA in a multi-cultural environment (see Q12).

Hence, a global TA study, for example on digitization issues, would not only be confronted with immense complexity but also with the absence of any powerful and mandated advisee. While for a national TA, the addressee is a main element in each exercise, a global TA may have to depart from the idea of a clear recipient. Instead, other channels of knowledge transfer have to be opened up. We will see, how this will develop in the future. Establishing global TA across different cultural contexts but also in different government and governance regimes is a significant challenge to our future work.

⑭ *GF: In your 2019 book, you discuss TA's obligations to transparency, inclusion, and democratic debate. How do you see the prospects for*

implementing TA in non-democratic political systems? Is it at all feasible, or do such contexts inherently limit its effectiveness and risk rendering it "doomed to fail"?

AG: TA is not a trademark, implying that everybody may use the term to describe anything. There is no formal reason preventing this. Please imagine, for example, a dictator who wants to install new surveillance technology in order to control and suppress the people more efficiently. He or she wants to be sure that the new technology will not have any unintended side effects for his/her dictatorship. In particular, the dictator postulates that that technology cannot be hacked or misused by his or her opponents. If he or she would commission a study to analyze the available technologies, investigate possible unintended consequences, compare them with regard to the objectives determined by the dictator and identify the most appropriate one – could this be regarded as TA or even as good TA based on sound TA methodology? Nobody could forbid doing so. However, my expectation is that declaring this activity as TA would be rejected by the global TA community.

The reason is that TA is built on a normative fundament rather than merely being a set of value-neutral instruments usable for arbitrary purposes. According to the three types of rationality following Rescher (see Q6) good TA is not only about cognitive and practical rationality but has also to observe evaluative rationality regarding the justifiability of the ends and purposes. At exactly this point, the dictator would fail attempting to declare his or her study as TA. Nobody would accept this (I hope).

In my conceptualization of TA, the dimension of *inclusion* is key in this respect. The underlying argumentation based primarily, in my reconstruction, on the political philosophy of John Dewey and Jürgen Habermas is among the major background narratives of TA in general. It is rooted in normative conceptions of humans and of the way humans should regulate their affairs. TA in a dictatorship would be a deficient version, lacking the conceptual dimension of *inclusion*, which must be a constitutive issue in the entire TA assessment process. TA is not value-neutral [63] and must not be used for purposes excluding the people affected. If these values were to be violated or if the procedures were to be undermined, e.g. when populist movements would go for discriminating and excluding particular groups (e.g. immigrants, refugees, people with a special sexual orientation, people following a minority religion) from access to civil rights, or if such movements want to abolish inclusive democracy, then TA cannot remain neutral, otherwise it would render itself inconsistent. In particular, in times of crisis of democracy TA has to take a stand, in order to defend and strengthen inclusion and democracy but also to develop them further. Probably, it is not accidental that most of the populist movements of the recent about fifteen years do not sympathize with TA.

The high emphasis on the commitment of TA to be inclusive and to support inclusion in this argumentation is not bound to inclusion in the sense of a Western democracy. If other cultures established other forms of inclusion, TA going to the global level beyond Western reasoning (see Q12) should be open and check whether these other types are in accordance with the very idea of inclusion. Even Western democracy is not a monolithic block. There are different understandings like deliberative democracy, grassroots democracy, representative democracy, liquid democracy etc. with different accentuations towards inclusion. Non-Western cultures and traditions could add new perspectives on inclusion. TA should be open in this respect.

⑮ *GF: Even within a single country, value pluralism poses a challenge: different stakeholders often hold divergent values and priorities. In this context, how do you approach the axiology of TA—particularly the balance between freedom, pluralism, and normative orientations? Are there specific TA approaches or methodologies that you consider better suited to navigating and integrating such value pluralism?*

AG: According to the development of TA debate over the recent decades [2, 34], there is no longer any doubt that TA is (1) based on values and (2) often is confronted with value conflicts in its assessments. The normative fundament for assessing technology futures, for shaping technologies, and for providing orientation for the digital and the sustainability transformation is decisive for assessments and decisions based on them. Above all, TA is bound to values of democracy since its origin, in particular in opposition to technocracy. Democratic values include, in particular, procedural justice, e.g. fairness and inclusion which are constitutive for TA (see Q26). Several TA approaches have been proposed and, at least partially,

been implemented explicitly referring to the corresponding frameworks of discourse ethics and deliberative democracy (see, [41, 42]).

While this type of values affects the manner how to organize TA assessments *procedurally*, other domains of values may lead the assessments *substantially*. In this respect, human rights [64] and the imperative of sustainable development are of utmost importance to TA [56]. However, what human rights or the *Leitbild* of sustainable development mean in a specific TA project, e.g. on gene editing, care robots or autonomous driving, is not easy to determine. Often, the consensus on observing both domains of values at a more abstract level gets fragile or even completely disappears when specific cases, technologies, and social contexts are considered.

Therefore, dealing with value conflicts due to pluralism belongs to the daily business of TA. This was an early observation already made in the 1970 s which, at that time, motivated parts of TA to remain cautious about engaging with value issues and, rather, to remain in the role of a neutral observer, just reporting about existing value conflicts. This more or less positivistic approach, however, has disappeared in favor of TA now explicitly dealing with values and value conflicts. In this sense, analyzing values and the relation between values as well as enlightening value conflicts in various, also normative respects belongs to TA. An early approach was the TA Guideline published in 1991 by the Association of German Engineers [53] which explicitly thematized value conflicts. It was co-designed by philosophers of technology and ethicists who were aware of the normative and evaluative dimension of assessments. Other TA approaches closely linked with ethics also took this direction early on, e.g. the Rational TA [65] and the Value Sensitive Design approach [66, 67]. The more recent RRI approach focuses on the value dimension, too (e.g. [49].

Dealing with value conflicts and pluralism, however, must not be done in TA beyond its own mandate of giving advice rather than predetermining decisions to be made by others. The distinction between TA as the advisor and, in case of policy-advising TA, the political advisees in charge of making legitimate decisions, is part of the basic TA configuration. Therefore, TA has to be careful to avoid becoming technocratic by determining solutions to value conflicts itself. At this point, it is time to recall the obligation of TA to support 'thinking in alternatives' instead of closing down the space of alternatives to a presumed 'one best solution' [68]. Balancing the explicit involvement with value conflicts and avoiding transgressing its own mandate require TA to be fully transparent in its assessment processes. Transparency in dealing with value conflicts allows for traceable critique of TA results in the sense of Social Epistemology [69] and offers added value to policy-makers and society by improving their understanding of value alternatives. This added value may also benefit participatory TA in order to support exploring solutions to value conflicts with various stakeholders. This approach allows keeping the role of the 'honest broker' [17] while simultaneously being evaluative [4].

⑯ *GF: In some cases, TA relies on funding from multiple sources, some of which may seek to influence the assessment process or its outcomes. How do you understand the relationship between these sources of power and the independence and autonomy of TA? How can TA maintain credibility and impartiality in the face of such pressures?*

AG: In order to put it clearly: the autonomy of TA is a *normative ideal* related to the role model of the 'honest broker', the approach of thinking in alternatives, the obligation to avoid technocratic or other biases but, rather, to follow principles of deliberative democracy. TA institutions and projects, however, do not operate in an ideal world but have to arrange with the many conditions and forces in the *real world*. The resulting tension is a specific manifestation of the general gap between the Ought and the Is. It is characteristic for TA's daily business but also applies to other approaches deemed to be independent but which are also depending on funds like Applied Ethics or Responsible Research and Innovation. It even applies to many other research activities made possible by third-party funding and being confronted with possible thematic expectations of the fundgivers. Hence, this is a rather general challenge in the science system.

An additional and more specific TA issue in this field originates from the tension between the independence resp. autonomy claimed by TA, on the one hand, and the situation that, on the other, TA is aimed at creating impact in the real world. For creating impact, it is important that TA is part of 'the system' (parliamentary democracy, liberal or

even neoliberal economy, national policies etc.). However, being part of the system makes it difficult, perhaps impossible, to criticize 'the system' or its basic convictions. Criticizing the dominant story of quantitative economic growth, for example, might cause resistance to TA rather than starting a society-wide reflection whether the goal of unlimited growth can be maintained in times of huge sustainability problems. In worst-case scenarios, this resistance could endanger the existence of TA-institutions and lower funds for TA projects. Being afraid of this could cause a self-censorship of TA resulting in a bias in favor of the 'existing system' and, e.g., contributing to "Sustaining the Unsustainable" [70] instead of openly criticizing the unsustainable parts of 'the system'.

These are serious issues in TA. However, they are not lethal as history has proven. The TA community and, in particular, TA institutions have learned to deal pragmatically with these tensions. Community and institutional mechanisms of reflection and quality assurance have been developed in order neither to run into the worst-case scenario mentioned nor to lose TA's own critical potential. In specific cases, institutional settings were organized in a way to tame the tensions mentioned. As an example, TA for the German *Bundestag* is, by regulation, bound to consensus among fractions represented in the Committee of Research, Education and Technology Assessment in charge of supervising TAB. The necessity to reach consensus for determining TAB's agenda, for commissioning TA projects as well as for approving the results ensures that TAB neither delivers socialist or conservative TA reports but always scientifically independent advice to the *entire Bundestag*. At the ITAS institute, the situation is different. ITAS receives, fortunately, considerable basic funding from the Helmholtz Association which is the largest research organization in Germany. This basic funding allows for a lot of freedom concerning the institute's agenda-setting and doing the research required. Furthermore, it is consensus in the institute that ITAS must not become dependent from private companies, e.g. by accepting huge projects funded by automobile or other companies.

Beyond political influence or lobbyism by large companies or associations, independence and autonomy of TA are also an issue in the cooperation with engineers [71] which applies primarily to TA approaches intimately linked with technical R&D and engineering like Value Sensitive Design. TA's independence might be threatened, if the necessary distance to the technical developments and those working on them would get lost, or if strong economic interests of project partners would diminish their willingness to accept independent assessments. If TA came up with unexpected and unfavorable results for the innovation under consideration, e.g. regarding the competitiveness or sustainability indicators, voices might come up to suppress these results, or to kick TA out of the project.

In all of these tensions and threats, the independence of TA is also an issue of the standing and reputation of TA researchers and institutes. In the more than 20 years I spent with ITAS and TAB so far, I experienced not extremely many but several situations of conflict, e.g. with other institute directors at KIT or with members of the *Bundestag*. As an experience I would like to say: first, mutual expectations and conditions must be clarified in advance, e.g. the opportunity to publish TA's results, and, second, it is better to reject a possibly problematic project at an early stage than accepting it and hoping to improve the situation over time.

In Germany, the shift from state dominance decades ago to the neoliberal system following "Thatcherism" and "Reagonomics" was not that hard like in many other Western countries. However, boundary conditions for innovation changed in face of increasing economic competition, in particular since the 1990s. Germany had years of economic crisis at the end of the preceding century. TA, therefore, was urged to put stronger emphasis on innovation. One example was renaming the former TA activities of the Federal Ministry of Research and Education (BMBF) into "Innovation and Technology Analyses" around the year 2000. The TAB soon integrated innovation studies into its portfolio of products offered to the Bundestag. And my "boss" in that time, the CEO of the Research Center Karlsruhe (a predecessor of KIT), recommended renaming the ITAS institute in "something with innovation". However, that was *Zeitgeist*. TA not only survived that time but started flourishing. Among the reasons was, I guess, the successful integration of innovation issues into its own

business. Interestingly, things went differently in other regions where parts of policy-advising TA moved to industry-advising Responsible Research and Innovation, aiming at filling the "legitimacy vacuum" [52] left over due to decreasing power of legitimacy-creating public institutions. This move is a clear consequence of the former rise of neoliberalism.

Part IV

Thinking about Technology, the Environment, and the More-than-Human World

⑰*GF: How do you see the relationship between Technology Assessment and environmental thinking, or more specifically, between TA and Environmental Ethics? In what ways can TA integrate ecological considerations and ethical perspectives on sustainability into its methods and practice?*

AG: The increasing awareness of environmental issues in the 1960 s and 1970 s was among the initial motivations of TA. Rachel Carson's "The Silent Spring" is usually classified a major root not only of the environmental movement but also of TA. Early TA was deeply interwoven with environmental problems because it quickly became obvious that they have something to do with technologies. At that time, it was mostly local or regional pollution of air and waters due to emissions of power plants and industrial production patterns. Global environmental challenges received public attention since the 1970 s only. The ozone hole and the political process to solve this problem (Montreal agreement 1982) were among the first typical global issues emerging from many sources distributed over industrialized countries across the globe. While the issue of unintended effects of the development, production, use, and disposal of technology to the environment was dominant in TA in its first decades, social, cultural and ethical issues gained larger significance in TA much later only.

In the 1970 s and 80 s, discussions about the legal establishment of environmental impact assessment (EIA) took place in many countries, partially accompanied by questions as to whether TA should be legally established, too. While the latter did not happen – except in the field of Health Technology Assessment HTA in some countries ([72]) – EIA or simply IA became mandatory in many countries and got formalized (see [73]). Frequently, life cycle approaches are used like the ISO-standardized Life Cycle Assessment (LCA).

In TA, LCA also plays a major role for measuring environmental impacts. There are, however, at least two important differences with far-ranging consequences. *First*, classical LCA compares existing products, services or objects with respect to their ecological footprints. In TA, however, *prospective* LCA [74] is needed for contributing to policy advice or to directly shaping technology according to environmental protection and sustainability goals. Dealing with prospective issues, however, is epistemologically different from collecting and exploiting present data.

Second, the targets differ. EIA was legally established as a regulatory instrument to be used in approval processes with possibly huge consequences for the permission to build e.g. industrial production plants, geothermal or wind power plants, or large buildings like bridges or dams. TA, however, is not about legal approvals but about providing orientation in situation of ambiguity, ambivalence, or conflict.

These differences have strong implications. EIA must be highly standardized and built on sound data to be robust with respect to possible opponents at court. TA, however, is and shall be much more explorative and open. It shall not just check the environmental situation but reflect on relations between technology, the environment, and society. In TA situations, sound data often are not available which could feed an EIA, because TA deals not with a present approval situation but with future issues.

As an example, and typical also for many other "new and emerging sciences and technologies" (NEST), almost no knowledge about possible impacts was available in the early times of nanotechnology around 20 or 25 years ago. Any EIA wouldn't have been feasible because of lack of data and the ignorance with respect to possibly emerging nano-based innovations. Classical risks assessment was not

doable, too. TA, however, was asked, e.g. by members of parliament in UK and Germany but also by ministries and departments, for orientation, and it had to respond. In this situation, TA developed approaches far away from the data-based LCA-type of reasoning, e.g. Vision Assessment, followed later by the approach of hermeneutic TA, but also experimenting with precautionary thinking (see, e.g., [75]). At that time, TA was exploring an unknown landscape of possible consequences, thereby preparing the ground for the situation in nanotech today [76], where EIAs of nanoproducts can be done in their established mode of operation.

Similarly, TA is again and again confronted with the challenge to investigate uncharted territory and to prepare the ground for subsequently needed and more standardized assessment procedures like EIA. At early stages of development TA may reflect about possible risk, possible benefit, and possible environmental consequences in order to *prepare* for coming debates. Sometimes, it is even not clear what may be regarded as benefit and what as burden or risk, and to whom these assignments would apply. Therefore, then it's more about *understanding* the possibly novel issues brought by new technology and their possible implications, not (yet) about specific assessments. This part of TA's business is rather far away from EIA or HTA and includes a strong hermeneutical component [77]. This relation applies to all fields in which TA has to deal with a future open to several alternatives, in particular in the field of new and emerging sciences and technologies (NEST) [78].

This is similar even for TA contributions to ongoing transformations, e.g. the mobility transition or the transformation of the energy system to a more sustainable, just, and climate friendly status. While standardized and mandatory EIA and IA shall directly support decision-making, e.g. in case of new highways to be built or the siting of a new waste disposal site, TA shall explore and structure unknown futures open to be shaped by current decisions. In this context, I like very much the following metaphor:

> … researchers, along with stakeholders, act as the "cartographers" of different, viable policy pathways and their practical consequences by acting as the "mapmakers" of the political solution space. They provide a guidebook with alternative options for policymakers (i.e. the "navigators" and the public). Such maps cannot replace travelling i.e., decision-making nor can they resolve all environmental policy conflicts, yet they can provide an important orientation in otherwise uncharted territory ([18], p. 63).

⑱ *GF: What is your "environmental moral position"?*

AG: Concerning this, my answer is not very specific: I don't have a pronounced position about the various approaches of biocentric etc. environmental ethics. In contrast to many present voices, I confess to be an anthropocentric. Before readers might feel provoked, I should explain it a little bit. Due to my conviction, humans are bound to methodological and epistemological anthropocentrism, not to an ontological or ethical one. We only can see the world including the environment through our glasses of recognition and epistemology as well as from our own perspective as being humans. We cannot exchange views and perspectives with animals, no discourse is possible with them about how to perceive the world. This diagnosis implies that we only have anthropomorphic perceptions of the environment: it is *our* perception, and we cannot escape from it. Even if we decide that animals or other members of the natural environments have rights (which is a good debate!), it is only us humans who make this decision, assign rights to them and then establish corresponding regulation. Even if we argue in a biocentric or pathocentric mode it's always us humans who think in this or that way and draw conclusions in a kind of advocatory care. In this sense, I'm a convinced anthropocentrist.

This methodological and epistemological self-understanding, however, does in no way predetermine any ethical or value statement about nature in general, animals, primates, or whatever. Reflecting on these issues is the next step only. In the context of the advocatory care mentioned above, human responsibility comes not only into the game but is, in my eyes, decisive. We humans – and that is because I like the contested notion of the Anthropocene – are, due to my knowledge the only species on Planet Earth which is able to reflect and take over planetary responsibility. I see our deep responsibility to take this much more seriously than humankind did over the last decades and centuries.

⑲ *GF: related, what is the role and status of thinking also about "non-human" or "more-than-human" entities and beings within TA?*

AG: Without any doubt, humans have been in the focus of TA since its origins, e.g. explicitly in techno-anthropological TA [79], in TA working for sustainable development applying the sustainability definition given by the Brundtland-Commission which is organized around *human* needs, or by investigating human/technology relationships in the ongoing digitalization and roboticization. In TA, humans are part of the game at several places and in various roles: technology is human-made, used by humans, has consequences for human life, changes human values and lifestyles, is deliberated and debated by humans, e.g. in participatory TA, and so on.

Non-human entities in TA are, on the one hand, technologies – I guess, this is not the target of your question. On the other, it is about non-human entities *affected* by technologies produced, used, and disposed by humans: animals, plants, landscapes, material flows, energy flows, the environmental media water, soil and air, ecosystems, and so on. These have been entities in TA in a passive sense since decades, affected, e.g., by unintended consequences of technology. Therefore, issues of biodiversity, the value of animals (see, e.g. the study on Animal Enhancement, [80] and others are part of TA practice.

However, I guess you have a further meaning of "non-humans" in mind. In the far-spread Actor/Network Theory (ANT) Bruno Latour postulated the shift from human-centered view on actors to including non-human objects as actors, too (actants). I concede that this is an interesting proposal for playing around with new perspectives and for breaking up traditional worldviews. However, as soon as it comes to deliberation and decision-making, humans are asked for. It is the same argument as given above (see Q18): we can try to see the world "as if" there were non-human actors. But we cannot invite them to join a participatory TA or to deliver their perspectives to members of parliament. We only can do this in an advocatory role – and then it's again us humans who represent the assumed perspective of non-human actors. Therefore, as soon as things get serious humans are needed. There is nobody else for, e.g. giving reasons, weighing arguments, making decisions, taking responsibility, and so on.

Part V

How to Think about TA? Who Thinks about TA?

⑳ *GF: In a 2010 article* [81] *that appeared in this journal, when it was still called NanoEthics, you wrote that:*

> "The Collingridge dilemma is exaggerated. The question of whether ethics should start early or late, whether it should be prospective or only start when reliable statements about consequences are available poses a false alternative. The issue here is not an either-or but a differentiation of ethical reflection in line with the problem at hand and with the validity of the available knowledge of the consequences. Ethical reflection differs conceptionally and methodically depending on whether it is concerned with measureable consequences of technology or with just imagined or even merely speculative ones. Above all, it also serves different purposes."

Then, you go on suggesting that, at least in the case at hand (i.e., nanotechnologies), it is better to move from "nanoethics" to "explorative philosophy of nanotechnology".

Would you say that this proposal may work more in general or is it specific to nanotech?

If it the latter (i.e., your idea of "explorative philosophy" is specific to nanotech), I would still like to ask, how would you propose to tackle the Collingridge dilemma for current "fast tech" developments beside nanotechs?

AG: The Control Dilemma [82] has indeed been among the most-referenced narratives in TA debates: in the early stages of technology development it is easy to shape the further R&D according to goals and values – but then reliable knowledge about later consequences of options at choice usually is not available. Better knowledge will be available in later stages – however, by then, path dependencies, e.g. with respect to costs and effort spent, will render attempts to modify the technology unfeasible.

An analogous debate was conducted on the ethics of technology (e.g., [83]), raising the concern that either ethics would be reduced to a "repair ethics" [84], always lagging behind and coming too late, or

being in risk to get purely speculative because being too early [85].

Looking at the TA practice, however, shows a different picture. While the Control Dilemma renders TA and ethics impossible, TA works in practice and does its job. How is this possible? The solution is looking at the presupposition of the Dilemma which is a strict dichotomy between early and late. This dichotomy is a theoretically constructed artifact. The question of TA being very early or only starting when reliable knowledge about consequences is available, poses a misleading alternative. It is never "too early" nor "too late" but always in between. Hence, differentiated TA reasoning must be tailored to the problem at hand and makes use of the respectively available knowledge. The question, for example, as to whether it is responsible to use nanoparticles to make food more durable is a concrete question to be considered in the context of regulation, consumer protection, labeling requirements, corporate commitments, and individual responsibility. Specific alternatives to support decision-making can be expected from a TA exercise. In contrast, early thoughts about synthetic biology or human enhancement serve the improvement of conceptual understanding, e.g. by developing a clear terminology, by understanding the novelty of the new opportunities proposed, by identifying the stakes involved and so forth [86]. The more speculative the considerations of the consequences, the less they can serve as orientation for concrete (political) action and decisions.

Consequently, we should not focus only on the sharp distinction provided by the Collingridge Dilemma but look for adequate and effective opportunities for TA to provide orientation through all the processes of research, development and innovation, hence accompanying these processes by enhancing reflexivity at their different stages. This view (see [4] is, with different nuances and accentuations, common to approaches like the Real-time TA [87], managing technology in society facing their co-evolution [24], and mid-term modulation [88]. Accordingly, the respective assessment processes have to apply different TA approaches and methods, according to the respective stage of development and knowledge available.

At early stages of development, reliable prospective knowledge will not be available and the consequentialist paradigm usually will not work. Here, one of the two branches of the Collingridge Dilemma fully applies. TA, however, is possible and may have impact. The term "explorative philosophy" mentioned in your question was coined exactly for this situation. E.g., hermeneutic TA [77] can be applied for improving the understanding of the respective debates and the hermeneutic circles on the societal meaning assigned to the technology under consideration [4], or analytic approaches such as vision assessment [89] can contribute to illuminating future visions and imaginaries disseminated and debated. Explorative philosophy in TA does not assess technologies but prepares the ground for later R&D stages. At mid-term stages, fields of application of the technology under consideration become better visible, knowledge about possible products and innovation paths becomes available and first considerations of the consequences and implications can be made. Then, e.g., scenarios for implementing innovation strategies and shaping appropriate boundary conditions for policy advice can be developed. In the later innovation phase, roadmaps and specific technology designs are under discussion. Then, constructive TA, systems analysis, value-sensitive design, or RRI can be applied in order to enrich ongoing engineering with prospective knowledge and ethical reflection. Hence, TA may enrich technology R&D as well as shaping innovation at different stages of development by applying different approaches and methods. The Collingridge Dilemma loses its scare and gets dissolved by overcoming the strict dichotomy between too early and too late. The music plays in between.

You recognized that I did not refer to nanotech so far. Indeed, while the idea of TA as an explorative philosophy at early stages was born on the occasion of the nanotech debate more than 15 years ago, it is not restricted to that field. In the meantime, we witnessed other fields of new and emerging science and technology (NEST) like human enhancement, synthetic biology, climate engineering, and care robotics which also were publicly debated at very early stages of development. Similarly to the early nanotech debate, again philosophical and hermeneutic questions dominated: what does human *enhancement* mean, how does human enhancement interfere with human nature, what is really novel about synthetic biology, how is synthetic biology related to reductionism and naturalism, what may be '*good* care' including robots, and so on. Therefore, the scheme

mentioned above was indeed developed on the occasion of nanotech but applies to all groundbreaking developments at the early stages of development.

You also ask what happens with the Collingridge Dilemma and TA in case of fast technology development. Acceleration as part of the capitalist economy unleashes creativity and fosters innovation through competition. Indeed, we are witnessing ongoing acceleration of many innovation processes due to digitalization and global competition. The increase in computing speed, the linking of creative resources via the Internet, and the acceleration of data transfer and communication shorten innovation cycles. This is usually appraised from an innovation policy point of view. However, TA aimed at 'strengthening reflexivity' [4] simply needs time. TA requires careful analysis and deliberation, the transparent weighing of alternatives, the search for the right balance and ethically legitimate criteria for choice-making for being able to do its job. Societal capacities and cognitive structures needed for reflection and TA could slowly erode. TA would be pulled to one of the two branches of the Collingridge Dilemma: at the end, it would always be too late, without any chance for enriching the innovation machinery by knowledge and reflection. This would be the ultimate victory of technology determinism.

Of course, this dystopian view on future TA is not a prediction. However, acceleration may indeed become ruinous, in this respect not only for TA but for any approach to shape technology with respect to social values and in accordance with the demands of stakeholders. Therefore, TA must keep an eye on the temporal structure of technology development, innovation, governance, and their dynamics – which is also decisive for the governance towards sustainable development [90].

㉑ *GF: What does it mean that TA is a scientific discipline? In which sense is TA "scientific"?*

AG: The claim that TA is "scientific" is a response to external expectations and demands. From its very beginning in the OTA, TA is expected to deliver the best scientific knowledge available on the respective issue for well-reflected assessments and orientation for action. In terms of philosophy, this is nothing else than TA's obligation to argumentative rationality and transparency in epistemic respect [4], Sect. 1.4), e.g. in the sense of Rescher [29] or Habermas [91]. The specific meaning of this obligation must be made operationalizable with respect to TA's quality criteria and ways how to check and prove their fulfillment. In this regard, I like referring to Social Epistemology [4]. Following Longino [69, 92], enabling the possibility of criticism is essential for ensuring quality. A culture of mutual criticism allows to check TA results from various perspectives and to make them more robust in moral, social but also scientific respects – or in other words: to increase the argumentative validity of TA statements. Criticism in TA must involve, according to its assessment process and the actor constellation [4]:

- criticism from scientific disciplines with respect to their respective bodies of knowledge, because any contradiction between TA results and established scientific knowledge obviously would constitute a serious problem;
- criticism from organized stakeholders with regard to their bodies of knowledge but also in consideration of their values and interests;
- criticism from citizens also interested in the issue under consideration but without having a direct stake – this is the epistemic dimension of participation (Sect. 3.3.1).

In order to enable these types of criticism TA has to take care of the *conditions of their possibility*. Realizing transparency is probably at the heart of enabling criticism and ensuring argumentative quality. Looking for the best argumentative validity, however, motivates the question of what is a good argument in TA. Good arguments are if/then-chains of reasoning [45] which may include scientific but also normative parts (as was an essential part of the rational TA approach mentioned above):

> Accepting the role of values in science does not eliminate the requirement for good arguments. It only modifies the understanding of what can count as a good argument [93], p. 560).

In TA, the disclosure of the "if"-compartments of any assessment result is of particular importance. Unclarified premises and presuppositions are a source of misunderstanding, ideology and dubious compromises; there is an emancipatory interest in making implicit assumptions explicit, in order to make them accessible to discussion and criticism [69] for improving the argumentative quality of TA results. Hence,

we arrive at a specific image of the results of TA and their scientific nature: they are complex sets of if/then-chains. The argumentative quality of TA results shows itself in the validity of the relations between the "if" and the "then" parts of the statements.

If TA is, in this sense, about structured if/then-chains, it is possible to consistently combine the thinking in alternatives as a crucial element of TA reasoning with the openness of the future and with the necessity for valid arguments. The openness of the future and the thinking in alternatives show themselves in the *antecentia*, in the "if" parts of the TA assessment results. The specific scientific expertise of TA is the creation and communication of those if/then-chains, while TA does not have a mandate to set the *antecedential* if-conditions into force. This is due to democratic deliberation and legitimated decision-making [45].

Following this argumentation, it seems quite obvious that technology assessment is a science- and research-based activity but not a science in the traditional understanding. Rather, TA seems to be an *art* – how to do things adequately, i.e., in coincidence with its cognitive interest, in socio-epistemic assessment practices with their own quality criteria [4]. Regarding TA as an art leads to the question for the artists: about skills and TA literacy, about capacity- and community-building, about an ethos of TA or perhaps a code of conduct, and about the specifics of an identified form of expertise.

㉒ *GF: Turning to the topic of religion and TA:*

Do you perceive any connections between the aims of institutional religions and the goals of Technology Assessment?

More broadly, do you see similarities, stark differences, or potential compatibilities or incompatibilities between theological or spiritual perspectives and approaches to technological development and innovation?

AG: To be honest, I never thought systematically about this question. Being a Christian faithful myself, it is and always was clear to me that TA must follow generalizable values like sustainability rather than particular belief systems from specific religions. This is, I guess, an intrinsic part of the ethos of TA. Correspondingly, I never included theology in my view on the many disciplines to be involved in TA. Also, no theologians are researchers at my institute.

While this is a correct but more formal consideration, things become more interesting when looking deeper. All the World religions emphasize, as far as I can see, human responsibility for creating a good world. It might differ what they understand by this, but they all have an eye on human responsibility among humans but also beyond, in particular including other forms of life. They know about human co-creatorship (in Christian terminology) concerning how Planet Earth looks like today but they also see the limited capability of humans who are regarded not being able to create paradise. Consequently, they postulate respect among humans and with regard to Nature while opposing ideas of human hubris and Francis Bacon's postulate of "full dominion over Nature".

In some sense, this is what also can be seen in modern history and technology. The technological progress has been accompanied by far-ranging expectations, sometimes reaching up to promises of creating paradise (e.g., in the early debate on nanotechnology). Transhumanists even dream of "paradise engineering". Reality, however, is radically different. Technological advances provided us without any doubt with a huge amount of fascinating new opportunities and benefits. I'm extremely happy about this progress, e.g. in the field of digitalization. However, did this fantastic development really bring us closer to paradise? While leaving this question open (personally, I have a clear answer), we clearly see that the wonderful technological advance has a 'dark side' – unintended and sometimes really dramatic or even devastating consequences. Technology and the technological advance, hence, show a deep and inherent ambivalence. Because the Anthropocene was made possible by science and technology only, this ambivalence also applies to our current era in total. It not only shows welfare, health, enhanced mobility, freedom, security and growth of consumption but simultaneously also climate change, microplastic pollution, devastating weapons with the capability of mass destruction, the Holocaust genocide with murdering millions of Jews and others with high technical efficiency, and so on.

In my mind, this history supports the religious conviction that we humans are not able to create paradise, that techno-optimistic hopes to overcome the historical experiences by ever more technological progress will fail, and that we should be more modest. Instead

of following fantastic hopes for technological salvation from all our problems, ways to a 'good Anthropocene' need, so my conviction, human responsibility and capability to navigate the further technological progress prudently, observing the *Leitbild* of Sustainable Development, instead of hoping that the next wave of miraculous technology will solve all of our problems and bring us closer to paradise [1].

In this respect, religions come into the game because of their similar motivation: acting on Planet Earth is not about creating paradise but about doing prudent next steps to lower harm and to improve life for humans and other species. For doing this, religious insights, wisdom, and spirituality may help to avoid running into the trap of naïve techno-optimism. In ethical respect, religions may support moral judgment on technology because many of the moral principles of religions do have a generalizable core. As an example, the Golden Rule of ethics is explicitly part of the Christian New Testament. This is not a partisan religious message valid only for Christians but, rather, covers a universal rule. As another example, the regulation of the use of water in Islam shows deep respect and also includes ideas of water as a common good which must not be reduced to a mere resource in economic respect. A lot of wisdom concerning our relations with Nature at large can, in particular, be found in Buddhism. The generalizable nature of many messages rooted in religion was demonstrated by the global and cross-culture perception and appreciation of the encyclical 'Laudato Si' by Pope Francis [94].

While there are many opportunities for synergies between religions and TA, there are also strict borderlines. For example, substantial positions of particular religions or churches with respect to the moral value of living organisms have to be checked for generalizability due to ethical standards of philosophy. The debate about the moral value of embryos demonstrates that there are severe controversies. Probably even more important is that the inclusive dimension of TA may collide with exclusive thinking and practices of specific religions, strictly separating between the faithful (in that specific sense) and "the others" which often are regarded having less value, being adversaries, having minor rights etc. This is not compatible with the normative fundament of TA. This fundament, in particular, includes democracy as a normative root of TA—which is in conflict, for example, with dictatorships and theocracies.

Hence, religious traditions, rules, wisdom, and spirituality form valuable resources for TA (and beyond). However, the messages and practices have to be checked for generalizability and for their compatibility with the secular normative fundament of TA.

㉓ *GF: What does it mean to be a human thinker in a robotic techno-sphere dominated by AI?*

AG: At first glance, the question sounds harmless. But behind the first glance, various deep and far-ranging questions are hidden. I'd like to start responding by giving my main message first: the response depends on our image of what humans are, and what AI-driven robots are. On the one hand, we can regard robots as technical artifacts, developed and produced as means to certain ends like a wheelchair or a washing machine. On the other, we can welcome robots as a kind of new species on planet Earth, new neighbors to us humans, artificial companions (analogously, they have artificial intelligence) which are not humans but very close already now, expecting that they will get closer and closer in the future. In this view, the close neighborhood has only one exception: in public perception, robots don't have human weaknesses and deficits, neither in physical respect but also in mental and ethical regards. They will become, as the prolongation of this perspective tells, the better humans. And indeed, then the sub-question of your question arises: will there be a place for humans in a future world dominated by AI and robots being much better than we humans?

The everyday way of talking about robots – and your questions breathes this attitude – is quite clear, supported in public perception by decades of consumption of science fiction literature and movies: Robots are not mere means to an end like traditional technology but are widely accepted as human-like beings. According to colloquial language, robots decide, learn, act, respond, show emotions – activities which we usually assign to humans. Some even postulate to assign human-like rights to robots. An explicit example is imagining the future production in industry ("industry 4.0"), where robots and humans are expected to cooperate "on equal footing". Questions remain, for example:

- What does it mean that humans and robots should cooperate "on equal footing"? Is symmetry in dis-

course meant, equal rights, participation in decisions, etc.?
- Can aspects of being human be determined in which the performance comparison with AI simply does not fit, is counterintuitive, or leads to arguments that are not comprehensible?
- Can aspects of being human be determined that cannot be replaced or improved by AI?
- What do possible AI equivalents of concepts like shame, guilt, or consolation look like?
- How do AI systems relate to the human ability to distinguish between the Is and the Ought?
- What does it mean anthropologically (if it means anything) that AI systems can show emotion? What does this mean, in turn, for the distinction between having emotions and showing emotions, or for the underlying question of the distinction between authentic and simulated emotions?

All these questions address the relationship between humans and AI systems. Ultimately, it is the renaissance of the old topic of who or what humans are, albeit here on a new foil of juxtaposition to AI systems as the creatures of humans. These observations call for an expansion of TA. While so far TA and the question for the human nature have hardly been mentioned in the same breath, the above observations lead in a different direction. The traditional anthropological abstinence of TA must be overcome in favor of cooperation between TA and philosophical anthropology. For the connection between TA and the question of the human cannot be separated from the question of our future, of hopes and fears, of limits and transgressions in the context of further technical progress.

The next issue concerns the phrase "dominated by AI" in your question. Can AI dominate anything by itself? AI consists of algorithms, software architectures, machine learning, and need huge data sets. Its strength is sophisticated mathematical statistics. AI calculates. AI does neither have a will nor intentions, it just calculates. Can AI being a fast calculation machine really become dominant in any respect? In this perspective, AI does not dominate anything – the dominance of AI and AI-driven robots is an *anthropomorphic assignment* of dominance. Dominance of AI means that we accept AI to be dominant. The effect happens on the side of humans: AI does not take over dominance but humans transfer their own capabilities of decision-making and acting to AI systems. They give up what Immanuel Kant put as "sapere aude" but simply trust the results of AI systems and assign dominance to them. In this perspective, we humans are not the victims of AI technology taking over control but we are guilty because we fail to shape the AI-influenced world into a humanist world.

Even worse, accepting the "dominance of AI" hides the fact that dominance is not exerted by "the AI" but by humans and corporative actors operating behind it and shaping its characteristics. AI is made by humans due to values, for business models, for security purposes, for opening up new markets, for becoming more competitive at the global level etc. This is another dimension of "dominance of AI": it is the dominance of those who were and are able to steer the development of AI, the major AI products and services, e.g. the Silicon Valley companies exploiting AI capabilities for their business purposes. If we blame AI technology and are afraid that it will take control over humans, we forget those standing behind the AI technology. This would be the ultimate victory of technology determinism which is the ideology of the powerful [4] because it allows to hide and forget about their responsibility.

㉔ *GF: Let me provoke you a bit more then: can and should AI systems or perhaps AI-powered robots do TA? If yes/no, why?*

AG: Yes, it seems that robots can do some kind of TA. AI can perform complex statistical calculations based on data and then write a proper text like the ChatGPT family does. Please test it: ask ChatGPT about consequences of gene editing, and you will probably get an answer beyond complete nonsense. But what does this mean, what does the ChatGPT answer mean, what does it mean that AI is capable to give this answer, and is this TA? Let's take the ChatGPT setting – we ask questions, AI responds – take as configuration for answering your question.

Having in mind what I responded to your many questions on my understanding of TA, the answer to the latter question must be: no, that is not TA. The reason for giving this answer is not simply that TA must be done by humans, this would be too easy. Instead, let's look at basics of TA, its conceptual dimensions of anticipation, inclusion, and complexity management. Can AI explore possible futures involving consequences of new technology? Being bound to calculate on the basis of data from the past (data from the future are not available, following e.g. Augustine of Hippo),

AI only can prolong patterns and trends discovered in old data. Humans, however, also can think into the future without or even against data from the past. Moreover, often AI is not useful when consequences of not yet existing technology are considered.

Next point: AI calculates a lot but does not include stakeholders, citizens, people affected etc. Instead, it responds to questions based on data. These data cannot include perspectives of e.g. stakeholders because they were not yet involved, perhaps don't even know anything about the technology under consideration and had no chance to be part of the assessment process. AI only can simulate such a process and hallucinate about results. Simulation does not replace real participation, deliberation, and dialogue. And the third one: complexity management based on decisions about relevance. AI has one major criterion of data relevance: statistical frequency among the data investigated, e.g. on the available Internet resources. For analyzing consequences of not yet existing technologies this will usually be inappropriate. For example, the frequently used Horizon Scanning for anticipatory purposes looks for weak signals among the data, not for strong signals with high frequency.

Finally, TA is bound to transparency due to epistemological as well as democratic reasons while AI systems are black boxes. Results of AI systems can neither be scrutinized nor traced back to the input data which is a fundamental requirement to TA. How should, in this situation, AI results on TA questions be assessed, e.g. for quality? Quality of TA is, to large parts, bound to the quality of its assessment process [4] – how can we TA's quality in absence of a traceable process? And who could or should do this? Another AI?

AI can only simulate TA but not do TA, as is similar in many other areas. Concerning TA, however, this is an existential difference: a simulation of TA is not TA because the traceable TA process is missing.

Having said this, however, we should look at AI in TA differently: neither for simulating TA nor for replacing TA by nicely looking and fast simulation but to use AI tools to support TA processes. The wonderful properties of AI should be exploited to improve data-based evidence of TA inputs wherever possible, to simulate possible futures, to learn from advanced statistical calculation, and so forth. This is, as I see, the future of AI in TA: welcomed new opportunities for the TA toolbox.

Part VI

Summing Up: A Backward Glance, a Forward Vision

㉕ *GF: Turning to the future of Technology Assessment, in the foreword to your 2019 book you write:* "My main underlying conviction is that TA is not merely a highly specialized endeavor at the interface of technology and society but rather an expression of a longer paradigm shift in the relation of science, technology, and society at large" (p. xviii).

Could you expand on this statement and explain what you see as the broader paradigm shift underlying TA?

Looking ahead, how do you envision the future of TA? In other words, will institutions like ITAS continue to be necessary, and if so, in what form or role?

AG: TA is an instrument to provide orientation at specific occasions but also mirrors an attitude which characterizes late modernity. It is the insight, or better the idea or even foreboding that we have a problem with fundamental issues of modernity like classical rational choice, quantitative growth, subtle persistent colonialism or neo-colonialism, technical efficiency and so on. Hence, taking care of a livable future seems to be much more than an issue of considering this or that technology or innovation by TA or RRI. Often, TA ends up with "big questions" or "grand topics". A TA study on care robotics for the German Bundestag, for example, raised the question for what society understands by 'good care', simultaneously to 'good life', far beyond aspects of the efficiency of the care system [95]. Similarly to RRI and sustainability research, the idea behind TA exceeds TA in its daily practice [96]. It is exactly that perhaps audacious claim which allowed me talking about a 'good' or at least a better Anthropocene above. TA involves some visionary perhaps utopian ideas that the world could be a better place, and that we humans have the potential to bring it a bit closer to that 'better place'. I'm convinced that TA without this background conviction would never have survived over 50 years and flourishes today. In this sense, TA transcends itself, not loudly and noisy by proclaiming itself as a kind of savior but in its background motivation in many debates.

Concerning the future of TA, a lot has already been said responding to previous questions. In an ideal world, humans perhaps wouldn't need TA

anymore. In that world, technology and innovation as well as their social embedding would be considered so carefully ex ante that no problematic effects would occur after their release. Everything would be fine, only positive effects, justly and fairly distributed … This is a nice idea but its realization is not very probable to put it cautiously. On the one hand, the vicious circle mentioned above is a severe obstacle. On the other, it is questionable whether this would really be a good idea. Strictly speaking, that idea is no more than old planning optimism, partly based on technology determinism (the technology dominates its consequences) and ignoring human creativity. Creativity and the openness of the future will (and shall, in my eyes!) persist, implying that the general TA configuration characterized by openness, ambivalences, and possible dialectic backlash will remain. Hence, also the demand for TA will remain. Perhaps, other names will be coined because introducing new terms often is an important part of progress in science and research regardless the novelties included as such. But also with other names, the basic challenge of shaping a good, or at least better future will remain a part of the *conditio humana*.

㉖ *GF: You announced that you will retire in 2027, congratulations! If you were to outline the trajectory of your own thinking about technological developments and innovation, how would you describe it? Could you reflect on the key experiences, influences, or turning points that have shaped your intellectual path?*

AG: Times have been changing, new challenges have emerged and I have had lots of experiences due to ITAS research, to developments in the TA community and beyond, as well as due to changing external requirements. In the 25+ years at ITAS and in the 33+ years working with and in TA, I could learn a lot. In order to be short and precise as most as possible, I'd like to focus on major developments in my thought which found their expression in new topics and interests, shifts and modified accentuations in looking at TA and beyond, extensions of previous approaches, etc.:

- directly after having entered the TA community in 1991, I became part of a scientific dispute between normative ethics (represented by Carl Friedrich Gethmann) and sociology (represented by, e.g. Gotthard Bechmann from ITAS and Johannes Weyer). This controversy helped a lot to sharply see the blind spots of each of the disciplines involved – which, thus, was an essential experience for later doing interdisciplinary work;
- in that context – I was, by occupation, on the side of normative ethics – participatory TA was more or less rejected. Even at that time, I did not follow this strict position. In the meantime, I became strongly interested in normative reasons, as given e.g. by American pragmatist philosopher John Dewey. In my Routledge book on TA [4], I even determined "inclusion" as one out of three conceptual and constitutive dimensions of TA;
- among my first publications in TA, the question of recognizability of consequences of technology, e.g. their predictability, already was a major issue. Having been skeptical about the possibility to predict already then, my skeptical attitude grew over time. This increasing skepticism, however, means that, unfortunately, prospective knowledge is highly uncertain'. Instead, I discovered the openness of the future, our ability to shape the future, as a value. For me, it is a good feeling to regard the future as (in many respects) an open space;
- technology determinism, however, was a strong position still in the 1990. At that time, I was happy to have discovered the cosmos of social constructivism of technology and Constructive TA, as well as the *Leitbild* Assessment, which allowed a conceptual move, from considering consequences of technologies to shaping technology. This extension led, e.g. to applying TA for shaping technologies towards more sustainability, and became a strong part of my understanding of TA. In this context I'm currently deeply concerned about the revival of old technology determinism in the shape of digital determinism [97]. Sometimes, I feel like an itinerant preacher;
- closely related, in the 1990s, the position that technology is value-neutral was still widespread, in particular among engineers and in the economy. Different from technology determinism, the development was more enjoyable: in my eyes, this position widely has disappeared;

- an important step in my thinking about engineering, technology, and TA was the extension of the established consequentialist view of TA (even included in its name in the German translation) by the hermeneutic approach. It took me quite a long time to oversee fully the consequences of this extension and to be able to formulate it in an at least somewhat clear manner. This slow process is apparent in my publications, from works on Vision Assessment since around 2005 to the development of the hermeneutical approach about ten years later;
- during these years, the temporality of technology, of reasoning, of modelling, of TA and other activities developed to a major background issue in my thinking. Recently, I coined terms like 'temporal circularity' and 'temporal asymmetry' [98] which shall make essential elements of TA a bit clearer;
- overall, the *Leitbild* of Sustainable Development developed to become a crucial element in my value hierarchy, comparable to democracy and human autonomy.

In spite of rising concerns – e.g. whether humankind will be able to cope with climate change before horrible situations could occur, or whether humankind will successfully circumvent the many temptations of digitalization – I kept my attitude to at least some optimism because I still think that we humans have great potential to act as *homo responsibilis* [96].

㉗ *GF: On a more personal note, as you near retirement after more than 25 years at ITAS and over 30 years shaping the field of Technology Assessment, what do you hope your legacy will be—both intellectually and for the people and institutions you have influenced?*

AG: First of all, seen from today it was a wonderful decision to choose TA as the mission of my professional life. I'm happy that I could contribute to TA at various levels, and I'm even happier that TA is now in a good position in several respects: with regard to institutions and networks worldwide, to theory and interrelations with the world of academia, to advising practices in a vast area of diverse fields and contexts, to community development and capacity-building, and probably even more. This was not always the case. For example, I did never forget the "earthquake" in the TA community after the closure of US American OTA in 1995,[4] and the German "earthquake" after closure of the TA Academy Baden Württemberg in 2002 which was the flagship TA institution in Germany that time.

I would like to wish that events of this type will never come again even if today the resilience of the TA community is much higher. Now, TA has arrived at various places where we have postulated to be for long times. Now, it's time to deliver. My personal legacy is not decisive. But because you asked, I should say a few words on this. Mostly, it is said that value conflicts, trade-offs, and pluralism are behind the diverging attitudes towards new technology in modern societies. Obviously, this is not wrong. However, I think this is only part of the story. The other part is the epistemology of prospective assessments, or the diverging attitudes towards the future between hope and fear, between high expectations and dystopian narratives, between optimism and pessimism, between the wish for certainty and the concern of determinism. Hence, TA often is about an inherently confusing mélange of knowledge, attitudes concerning the future, and values. It's not just consequentialism, investigating consequences of technology and then assessing what we want to have and what not. Rational TA – an approach to which I contributed and still benefit in some respect – underestimated the significance of this mélange.

I became aware of it on the occasion of two distinct developments: the climate change issue which was trapped between prediction and precaution about 30 years ago, and the nanotech debate where established TA consequentialism simply did not work. Seen from today, the storyline of bringing the hermeneutic view to TA, first on the occasion of visions, but then extended, recently even to the field of model-based scenario-making [98], began there. I'm convinced that hermeneutics is needed to shed more light on the mélange mentioned above, in order to avoid bias, to realize transparency, to be inclusive and democratic. I hope that this insight will accompany future TA and will be further explored and made fruitful for good TA, for a good Anthropocene.

[4] To be honest, at that time I decided to leave TA, and I really left in late 1995 – fortunately, I was fetched by Carl Friedrich Gethmann few months later only at the occasion of the foundation of the European Academy Bad Neuenahr-Ahrweiler dedicated to TA and science ethics.

Funding Not applicable.

Data availability Not applicable.

Declarations

Ethics approval and consent to participate Not applicable.

Competing interests The two authors, in this case interviewer and interviewee work in different roles at the same institute.

References

1. Grunwald A, Krings BJ, Lösch A, Scheer D. (Eds.) (2025) Back to the roots: The conceptual development of TA in Germany. Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783748963073
2. Grunwald A (1999) Technology assessment or ethics of technology? Ethical Perspect 2:170–182. https://doi.org/10.2143/EP.6.2.505355
3. Grunwald A (2025) Technology assessment – Between sociological systems theory and philosophical ethics. In A. Grunwald, B.-J. Kring, A. Lösch, & D. Scheer (Eds.), Back to the roots: The conceptual development of TA in Germany (pp. 371–388). Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783748963073
4. Grunwald A (2019) Technology assessment in practice and theory. Routledge. https://www.routledge.com/Technology-Assessment-in-Practice-and-Theory-1st-Edition/Grunwald/p/book/9781138337084
5. Mitcham C (1994) Thinking through technology: The path between engineering and philosophy. The University of Chicago Press, Chicago
6. Funtowicz SO, Ravetz JR (1993) Science for the post-normal age. Futures 25(7):739–755. https://doi.org/10.1016/0016-3287(93)90022-L
7. Gibbons M, Limoges C, Nowotny H, Schwartzman S, Scott P, Trow MA (2012) The new production of knowledge: The dynamics of science and research in contemporary societies (1st ed). SAGE Publications
8. Wynne B (1975) The rhetoric of consensus politics: A critical review of technology assessment. Res Policy 4(2):108–158. https://doi.org/10.1016/0048-7333(75)90028-1
9. Van Den Hoven J, Vermaas PE, Van De Poel I (Eds.) (2015) Handbook of ethics, values, and technological design: Sources, theory, values and application domains. Springer Netherlands. https://doi.org/10.1007/978-94-007-6970-0
10. Laas K, Taylor S, Miller CZ, Brey EM, Hildt E (2022) Views on ethical issues in research labs: A university-wide survey. Account Res 29(3):178–201. https://doi.org/10.1080/08989621.2021.1910503
11. Jochem E (1975) Möglichkeiten und Grenzen der Technikfolgen-Abschätzung und Bewertung (TA), dargestellt an einigen ihrer Grenzen. Haas, H.(Hg.): Technikfolgenabschätzung. Köln, 55–66
12. Lee RP, Poganietz WR (2024) Beyond systems analysis to a multidimensional approach in technology assessment. In A. Grunwald (Ed.), Handbook of technology assessment (pp. 332–342). Edward Elgar Publishing. https://doi.org/10.4337/9781035310685.00048
13. Schneidewind U, Singer-Brodowski M, Augenstein K, Stelzer F (2016) Pledge for a transformative science: a conceptual framework. (Wuppertal Paper, 191). Wuppertal: Wuppertal Institut für Klima, Umwelt, Energie gGmbH. https://nbn-resolving.org/urn:nbn:de:bsz:wup4-opus-64142
14. Jonas H (1984) The imperative of responsibility: In search of an ethics for the technological age (1st ed). University of Chicago Press, Chicago
15. Leyten AJM, Smits REHM (1991) Technology assessment: Waakhond of speurhond?: Naar een Integraal Technologiebeleid. Kerckebosch
16. Dobroć P, Krings B-J, Schneider C, Wulf N (2018) Alternativen als Programm: Plädoyer für einen Perspektivenwechsel in der Technikfolgenabschätzung. TATuP - Zeitschrift Für Technikfolgenabschätzung in Theorie Und Praxis 27(1):28–33. https://doi.org/10.14512/tatup.27.1.28
17. Pielke RA (2010) The honest broker: Making sense of science in policy and politics (5. print). Cambridge University Press, Cambridge
18. Edenhofer O, Kowarsch M (2015) Cartography of pathways: A new model for environmental policy assessments. Environ Sci Policy 51:56–64. https://doi.org/10.1016/j.envsci.2015.03.017
19. Grunwald A (2017) Assigning meaning to NEST by technology futures: Extended responsibility of technology assessment in RRI. J Respons Innov 4(2):100–117. https://doi.org/10.1080/23299460.2017.1360719
20. Decker M, Ladikas M (2004) Technology assessment in Europe; Between method and impact—The TAMI project. In M. Decker, M. Ladikas, S. Stephan, & F. Wütscher (Eds.), Bridges between science, society and policy (pp. 1–10). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-06171-8_1

21. Petermann T, Bradke H, Lüllmann A, Poetzsch M, Riehm U (2011) What happens during a blackout: Consequences of a prolonged and wide-ranging power outage (Vol. 4). Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag (TAB)
22. Paschen H (1975) Technology Assessment als partizipatorischer und argumentativer Prozess. Heinz Haas (Hg.): Technikfolgen-Abschätzung. München: Oldenbourg Wissenschaftsverlag. S, 45–54
23. Böschen S, Dewald U (2018) TA als Kontextualisierungsexpertise: Zwischen einfachem und reflexivem Modus. TATuP - Zeitschrift Für Technikfolgenabschätzung in Theorie Und Praxis 27(1):34–39. https://doi.org/10.14512/tatup.27.1.34
24. Rip A, Misa TJ, Schot J (Eds.) (1995) Managing technology in society. Pinter Publishers, London
25. Grin J, Van De Graaf H, Hoppe R (1997) Technology assessment through interaction: A guide. Rathenau Institute
26. Joss S. (Ed.) (2002) Participatory technology assessment: European perspectives. Center for the Study of Democracy
27. Grunwald A (2025) Technikfolgenabschätzung. Einführung. 4. erweiterte und aktualisierte Aufl., Baden-Baden: NOMOS
28. Dewey J (1988) Human nature and conduct (Vol. 14). Carbondale: Southern Illinois University Press. (Originally published by Henry Holt and Co., New York, 1922)
29. Rescher N (1988) Rationality: A philosophical inquiry into the nature and the rationale of reason. Clarendon Press, Oxford
30. Gethmann CF (2023) Konstruktive Ethik: Einführung und Grundlegung, vol 52. Springer. https://doi.org/10.1007/978-3-662-66671-5
31. Habermas J (1985) The theory of communicative action: Volume 1: Reason and the rationalization of society (Vol. 1). Beacon Press, Boston. Translated by Thomas McCarthy
32. Habermas J, Lenhardt C, Habermas J (1991) Moral consciousness and communicative action (2. print). MIT Press, Cambridge
33. Grunwald A (2022) Parlamentarische Technikfolgenabschätzung und Risikokommunikation. Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz 65(5):552–558. https://doi.org/10.1007/s00103-022-03523-0
34. Nierling L, Torgersen H (Eds.) (2020) Die neutrale Normativität der Technikfolgenabschätzung: Konzeptionelle Auseinandersetzung und praktischer Umgang (1. Auflage). Nomos
35. Carpenter SR (1977) Philosophical issues in technology assessment. Philos Sci 44(4):574–593. https://doi.org/10.1086/288769
36. Bijker WE, Hughes TP, Pinch T (Eds) (1987) The Social construction of technological systems. Papers of a workshop held at the University of Twente, The Netherlands, in July 1984. The MIT Press, Cambridge, Massachusetts, London, England
37. Grunwald A (1996) Die Bewältigung von Technikkonflikten. Theoretische Möglichkeit und praktische Relevanz einer Ethik der Technik in der Moderne' in Zeitschrift für philosophische Forschung 51(1996)3:437–452
38. Petermann TH (1991) (Hg.), Technikfolgen-Abschätzung als Technikforschung und Politikberatung. Frankfurt
39. Decker M, Wütscher F (Eds.) (2001) Interdisciplinarity in technology assessment. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-04371-4
40. Moore GE (1905) The nature and reality of objects of perception. Proc Aristotelian Soc 6:68–127
41. Renn O (with Webler, T., & Wiedemann, P. M.) (1995) Fairness and competence in citizen participation: Evaluating models for environmental discourse. Springer Netherlands
42. Skorupinski B, Ott K (2002) Technology assessment and ethics. Poiesis praxis: Int J Technol Assess Ethics Sci 1(2):95–122. https://doi.org/10.1007/s102020100010
43. Schwemmer O (1987) Handlung und Struktur: Zur Wissenschaftstheorie der Kulturwissenschaften (1. Aufl). Suhrkamp, Frankfurt am Main
44. The Hegeler Institute (1930) The public and its problems: Monist, 40(4):640–640. https://doi.org/10.5840/monist193040413
45. Grunwald A (2003) Technology assessment at the German bundestag: 'Expertising' democracy for 'democratising' expertise. Sci Public Policy 30(3):193–198. https://doi.org/10.3152/147154303781780515
46. Harremoës P, Von Krauss MK (2002) Caring, daring and precaution. SETAC 23rd Annual Meeting Abstract Book, 69–69. https://orbit.dtu.dk/files/6509653/MR2002-148.pdf
47. Munthe C (2011) The price of precaution and the ethics of risk, vol Vol. 6. Springer, Netherlands. https://doi.org/10.1007/978-94-007-1330-7
48. Grunwald A (2011) Responsible innovation: bringing together technology assessment, applied ethics, and STS research. Enterp Work Innov stud 7:9–31
49. Grinbaum A, Groves C (2013). What is "Responsible" about responsible innovation? Understanding the ethical issues. In R. Owen, J. Bessant, & M. Heintz (Eds.), Responsible Innovation (1st ed., pp. 119–142). Wiley. https://doi.org/10.1002/9781118551424.ch7
50. Stilgoe J, Owen R, Macnaghten P (2013) Developing a framework for responsible innovation. Res Policy 42(9):1568–1580. https://doi.org/10.1016/j.respol.2013.05.008
51. Blok V, Lemmens P (2015) The emerging concept of responsible innovation. Three reasons why it is questionable and calls for a radical transformation of the concept of innovation. In B.-J. Koops, I. Oosterlaken, H. Romijn, T. Swierstra, & J. Van Den Hoven (Eds.), Responsible innovation 2 (pp. 19–35). Springer International Publishing. https://doi.org/10.1007/978-3-319-17308-5_2
52. Pellizzoni L (2004) Responsibility and environmental governance. Environ Polit 13(3):541–565. https://doi.org/10.1080/0964401042000229034
53. VDI-Gesellschaft (Ed.) (1991) Auftragsabwicklung im Maschinen- und Anlagebau. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-95783-3
54. Decker M, Grunwald A (2001) Rational technology assessment as interdisciplinary research. In M. Decker & F. Wütscher (Eds.), Interdisciplinarity in technology

assessment: Implementation and its chances and limits (pp. 33–60). Springer. https://doi.org/10.1007/978-3-662-04371-4_4
55. Grunwald A, Saretzki T (2020) Demokratie und Technikfolgenabschätzung: Praktische Herausforderungen und konzeptionelle Konsequenzen. TATuP - Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis 29(3):11–17. https://doi.org/10.14512/tatup.29.3.11
56. Böschen S, Grunwald A, Krings BJ, Rösch C (Eds.) (2021) Technikfolgenabschätzung: Handbuch für Wissenschaft und Praxis. Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783748901990
57. Gee D, Greenberg M (2013) Asbestos: From 'magic' to malevolent mineral. In: Paul Harremoes, David Gee, Malcom MacGarvin, Andy Stirling, Jane Keys, Brian Wynne, Sofia Guedes Vaz (ed) The precautionary principle in the 20th century: Late lessons from early warnings, 1st Edition, Routledge, pp. 49–63. https://doi.org/10.4324/9781315071985
58. Beck U (1986) Risikogesellschaft: Auf dem Weg in eine andere Moderne. Suhrkamp
59. Beck U, Giddens A, Lash S (1994) Reflexive modernization: Politics, tradition and aesthetics in the modern social order. Stanford University Press
60. Hahn J (2019) Towards a global technology assessment—Insights from cases in Germany, China, India and beyond [PhD Thesis, Karlsruher Institut für Technologie (KIT)]. https://doi.org/10.5445/IR/1000093530
61. Hennen L, Hahn J, Ladikas M, Lindner R, Peissl W, van Est R (Eds.) (2023) Technology assessment in a globalized world: Facing the challenges of transnational technology governance. Springer International Publishing. https://doi.org/10.1007/978-3-031-10617-0
62. Hennen L, Nierling L (2012) A new landscape of TA in Europe? First comparative findings from the PACITA explorative country studies 2012. International Workshop on Expanding the TA-Landscape in Europe, Karlsruher Institut für Technologie (KIT), November 15-16, 2012. https://publikationen.bibliothek.kit.edu/230089877
63. Delvenne P, Parotte C (2019) Breaking the myth of neutrality: Technology assessment has politics, technology assessment as politics. Technol Forecast Soc Change 139:64–72. https://doi.org/10.1016/j.techfore.2018.06.026
64. Ehrensperger E (2021) 5.1 Technologiefolgen-Abschätzung und die Menschenrechte als Motor der Demokratieentwicklung. In S. Böschen, A. Grunwald, B.-J. Krings, & C. Rösch (Eds.), Technikfolgenabschätzung (pp. 389–402). Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783748901990-389
65. Gethmann CF (2025) Rational technology assessment1. In A. Grunwald, B.-J. Krings, A. Lösch, & D. Scheer (Eds.), Back to the roots: The conceptual development of TA in Germany (pp. 361–370). Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783748963073
66. Friedman B (1996) Value sensitive design. INeractions 3(6):16–23
67. Friedman B, Kahn PH, Borning A, Huldtgren A (2013) Value sensitive design and information systems. In N. Doorn, D. Schuurbiers, I. Van De Poel, & M. E. Gorman (Eds.), Early engagement and new technologies: Opening up the laboratory (Vol. 16, pp. 55–95). Springer Netherlands. https://doi.org/10.1007/978-94-007-7844-3_4
68. Stirling A (2008) "Opening Up" and "Closing Down": Power, participation, and pluralism in the social appraisal of technology. Sci Technol Hum Values 33(2):262–294. https://doi.org/10.1177/0162243907311265
69. Longino HE (1990) Science as social knowledge: Values and objectivity in scientific inquiry. Princeton University Press, Princeton
70. Blühdorn I (2007) Sustaining the unsustainable: Symbolic politics and the politics of simulation. Environ Polit 16(2):251–275. https://doi.org/10.1080/09644010701211759
71. Grunwald A, Achternbosch M (2013) Technology assessment and approaches to early engagement. In N. Doorn, D. Schuurbiers, I. Van De Poel, & M. E. Gorman (Eds.), Early engagement and new technologies: Opening up the laboratory (Vol. 16, pp. 15–34). Springer Netherlands. https://doi.org/10.1007/978-94-007-7844-3_2
72. Banta D (2009) What is technology assessment? Int J Technol Assess Health Care 25(S1):7–9. https://doi.org/10.1017/S0266462309090333
73. Bond A, Dusik J, Ladikas M (2024) Impact assessment versus technology assessment: Distant relatives or different species? In: A. Grunwald (Ed.), Handbook of technology assessment (pp. 418–429). Edward Elgar Publishing. https://doi.org/10.4337/9781035310685.00058
74. Erakca M, Baumann M, Helbig C, Weil M (2024) Systematic review of scale-up methods for prospective life cycle assessment of emerging technologies. J Clean Prod 451:142161. https://doi.org/10.1016/j.jclepro.2024.142161
75. Grunwald A (2009) Nanoparticles: Risk management and the precautionary principle. In Jotterand, F. [Hrsg.] Emerging conceptual, ethical and policy issues in bionanotechnology Dordrecht: Springer Science and Business Media, 2008 (Philosophy and Medicine ; 101) Also publ. online (pp. 85–102)
76. Grunwald A (2010) Technikfolgenabschätzung - eine Einführung: Zweite, Grundlegend überarbeitete und Wesentlich Erweiterte Auflage, 1st edn. Nomos Verlagsgesellschaft mbH & Co, KG
77. Mehnert W, Grunwald A (2024) Hermeneutic technology assessment. In A. Grunwald (Ed.), Handbook of Technology Assessment (pp. 281–290). Edward Elgar Publishing. https://doi.org/10.4337/9781035310685.00041
78. Grunwald A (2014) Technikfolgenabschätzung als „Assessment" von Debatten. TATuP - Zeitschrift Für Technikfolgenabschätzung in Theorie Und Praxis 23(2):9–15. https://doi.org/10.14512/tatup.23.2.9
79. Botin L, Børsen TH (Eds.) (2021) Technology Assessment in Techno-Anthropological Perspective. (OA ed.) Aalborg Universitetsforlag. Serie om lærings-, forandrings- og organisationsudviklingsprocesser Vol. 10
80. Ferrari A, Coenen C, Grunwald A, Sauter A (2010) Animal Enhancement. Neue technische Möglichkeiten und ethische Fragen. Bundesamt für Bauten und Logistik. https://www.ekah.admin.ch/inhalte/ekah-dateien/dokumentation/publikationen/EKAH_Animal_Enhancement_Inh_web_V19822.pdf

81. Grunwald A (2010) From speculative nanoethics to explorative philosophy of nanotechnology. NanoEthics 4(2):91–101. https://doi.org/10.1007/s11569-010-0088-5
82. Collingridge D (1982) The social control of technology. Martin's press, Frances Pinter St
83. Moor J, Weckert J (2004) Nanoethics: Assessing the nanoscalenanoethics from an ethical point of view. In Baird, Davis, A. Nordman, & J. Schummer (Eds.), Discovering the nanoscale (pp. 301-). IOS Press. https://books.google.com/books?hl=en&lr=&id=O-LM6eWUbmIC&oi=fnd&pg=PA301&dq=Nanoethics:+Assessing+the+Nanoscale+from+an+Ethical+Point+of+View&ots=mgeLXR-IsC&sig=KmnOOog-yPz0y2yHvm0tH-B-acw
84. Mittelstrass J, Allgemeine Gesellschaft für Philosophie in Deutschland (1999) Die Zukunft des Wissens: XVIII. Deutscher Kongress für Philosophie, Konstanz 1999 ; Workshop-Beiträge. Northwestern University, UVK, pp 1362. ISBN: 3879406979, 9783879406975
85. Nordmann A (2007) If and then: A critique of speculative nanoethics. NanoEthics 1(1):31–46. https://doi.org/10.1007/s11569-007-0007-6
86. Grunwald A, Hartlieb J von (2012) Ist Technik die Zukunft der menschlichen Natur ? 36 Essays. Wehrhahn
87. Guston DH, Sarewitz D (2002) Real-time technology assessment. Technol Soc 24(1–2):93–109. https://doi.org/10.1016/S0160-791X(01)00047-1
88. Fisher E, Mahajan RL, Mitcham C (2006) Midstream modulation of technology: Governance from within. Bull Sci Technol Soc 26(6):485–496. https://doi.org/10.1177/0270467606295402
89. Schneider C, Rosmann M, Losch A, Grunwald A (2023) Transformative vision assessment and 3-D printing futures: A new approach of technology assessment to address grand societal challenges. IEEE Trans Eng Manag 70(3):1089–1098. https://doi.org/10.1109/TEM.2021.3129834
90. Bornemann B, Strassheim H (2019) Governing time for sustainability: Analyzing the temporal implications of sustainability governance. Sustain Sci 14(4):1001–1013. https://doi.org/10.1007/s11625-019-00683-y
91. Habermas J, Ben-Habib S (1981) Modernity versus postmodernity. New German Critique 22:3. https://doi.org/10.2307/487859
92. Longino HE (2013) Studying human behavior: How scientists investigate aggression and sexuality. The University of Chicago Press, Chicago
93. Douglas H (2000) Inductive risk and values in science. Philos Sci 67(4):559–579. https://doi.org/10.1086/392855
94. Grunwald A (2017b) Prophetische Worte - die Enzyklika "Laudato Si'. Talk at Katholische Hochschulgemeinde, Karlsruhe, Germany, 17 January 2017
95. Kehl C (2018) Robotik und assistive Neurotechnologien in der Pflege—Gesellschaftliche Herausforderungen
96. Grunwald A (2021) Living technology. Philosophy and ethics at the crossroads between life and technology. Jenny Stanford
97. Mainzer K (2016) Toward a theory of intelligent complex systems: From symbolic AI to embodied and evolutionary AI. In V. C. Müller (Ed.), Fundamental issues of Artificial Intelligence (Vol. 376, pp. 241–259). Springer International Publishing. https://doi.org/10.1007/978-3-319-26485-1_15
98. Grunwald A (2023) The hermeneutic perspective on modeling in technology assessment. In A. Grunwald, A. Nordmann, & M. Sand, Hermeneutics, History, and Technology (1st ed., pp. 172–192). Routledge. https://doi.org/10.4324/9781003322290-13