Technology assessment and the governance of automated vehicles: a Collingridge-dilemma or a lack in normative orientation?

Jens Schippl

To cite this article: Jens Schippl (2024) Technology assessment and the governance of automated vehicles: a Collingridge-dilemma or a lack in normative orientation?, Journal of Responsible Innovation, 11:1, 2342594, DOI: 10.1080/23299460.2024.2342594

To link to this article: https://doi.org/10.1080/23299460.2024.2342594

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 30 Apr 2024.

Submit your article to this journal

Article views: 76

View related articles

View Crossmark data
Technology assessment and the governance of automated vehicles: a Collingridge-dilemma or a lack in normative orientation?

Jens Schippl

Institute for Technology Assessment and Systems Analysis (ITAS), Karlsruhe Institute of Technology, Karlsruhe, Germany

ABSTRACT

It is often assumed, at least implicitly, that responsible governance of automated vehicles (AVs) requires more knowledge about the future development of the innovation and its potential consequences. In this context, technology assessment (TA) studies often refer to the so-called Collingridge-dilemma. This paper argues that, at least in the German case, a lack of knowledge in the sense of the Collingridge-dilemma is not the central challenge for the governance of AVs. The argument is developed on the basis of different types of knowledge for TA recently introduced by Armin Grunwald. The paper shows that responsible governance of AVs requires more normative and hermeneutic knowledge to better understand the directionality of the current system. More important than focusing on the possible consequences of AVs is a better understanding of how to overcome existing obstacles to the development of a broadly shared vision with effective goals for the German mobility sector.

ARTICLE HISTORY

Received 11 May 2023
Accepted 9 April 2024

KEYWORDS

Technology assessment; automated vehicles; Collingridge-dilemma; directionality

Introduction

Automated vehicles (AVs) are being tested in many pilot projects in different regions of the world. In the US, companies such as Waymo, Zoox or Cruise in Phoenix or San Francisco show that driverless vehicles can already operate in urban traffic under certain conditions (Yen et al. 2024). In Germany, a law was recently passed to allow highly automated driving under certain operating conditions. Many experts believe that automation in road transport will continue to advance (Lyons 2022). However, how this process will unfold is still an open question. Science and Technology Studies (STS) scholars have repeatedly emphasised that AVs technology in any form should not be taken for granted, but that alternative futures of mobility with and without technological progress should be considered (Stilgoe and Mladenović 2022; Van Wynsberghe and Pereira 2022). Although it remains controversial whether and when highly
automated driving, or even vehicles without a steering wheel, will be a common feature of the roads, there is widespread agreement that the new technology has very high transformative potential and can be of great importance for the future of the entire mobility sector (Cohen and Cavoli 2019; Milakis, van Arem, and van Wee 2017). Self-driving vehicles could fundamentally change the routines of road users and open up new perspectives: new daily routines, new travel habits, new business models, new urban structures etc. In view of this far-reaching potential for change, it is important to ensure that the development of AVs is in line with the ideas of a sustainable and climate-friendly development of the mobility system (Hopkins and Schwanen 2018; Stilgoe 2018). However, this does not seem at all certain. Discussions and publications in recent years show that AVs can affect the mobility system, infrastructures and also spatial structures in very different directions (Legène et al. 2020). In principle, both a significantly more sustainable and a significantly less sustainable development are conceivable.

Given the great transformative potential that is often attributed to AVs, it is important to anticipate developments at an early stage and to guide them in a desired direction in good time. However, this seems to be hampered by a problem that is characteristic of many topics in the field of technology assessment (TA) and which is expressed in the oft-cited Collingridge-dilemma (Collingridge 1980; Genus and Stirling 2018). AVs are a new technology that is not yet established, and its development dynamics and impacts are subject to many uncertainties. As approaches from technology assessment and transition research have long revealed, the mobility system must be understood as a sociotechnical system (Fraedrich, Beiker, and Lenz 2015; Geels et al. 2012; Ryghaug et al. 2023). Technical and social dynamics influence each other, often leading to path dependencies or lock-ins that are difficult to correct or even reverse. Such co-evolutionary processes are complex and fundamentally difficult to anticipate (Markard, Raven, and Truffer 2012; Truffer, Schippl, and Fleischer 2017). Nevertheless, planning must be possible in order to fulfil public welfare tasks and to be able to avoid a socially undesirable development of AVs and mobility systems. For these reasons, this field of technology has become a core topic of technology assessment in the mobility sector (Milakis, van Arem, and van Wee 2017; Truffer, Schippl, and Fleischer 2017).

This paper investigates the extent to which the challenges described in the Collingridge-dilemma actually hinder or even prevent effective governance of AVs. The focus is on Germany, a country where the automotive sector has traditionally played a key role. The argumentation is based on approaches from TA and transition research (section 2). With regard to transition research, particular reference is made to the concept of ‘directionality failures’ (Weber and Rohracher 2012) and the related phenomenon of suboptimal path dependencies in sociotechnical systems. With regard to TA, the argumentation takes the Collingridge-dilemma as the starting point of the problem and then draws on two typologies introduced by Armin Grunwald: the classification of five different ways of epistemological insight or types of knowledge relevant for TA (Grunwald 2020) and the distinction of three modes of TA (Grunwald 2014). Section 3 intends to show that for a sustainable development of AVs and the mobility system, critical deficits exist mainly in the domains of normative knowledge and hermeneutic knowledge. Section 4 illustrates that these deficits in the area of normative knowledge are both characteristic and problematic for ambivalent technologies such as AVs, which correspond to what will be described as ‘Mode-2-directionality’, because here path
dependencies can arise that can be classified as socially undesirable from several points of view. Conclusions are drawn in section 5.

**Conceptual background**

**The initial problem: AVs and the Collingridge-dilemma**

Dealing with new technologies is usually a knowledge problem and a time problem. The two are closely related. Knowledge about future developments and consequences of a technology accumulates over time and thus improves, but complete certainty about future consequences is never achievable, especially since not only the technology may evolve but also the context may change. Thus, the question can always be asked at what level of knowledge or uncertainty what kind of action may become legitimate or even necessary.

This issue of timing, and the challenges it poses for the social ‘control’ of technological developments, is addressed in the Collingridge-dilemma (Collingridge 1980). The first horn of the dilemma is that in the early stages of technological development there is an information problem; as long as a new technology is poorly developed and hardly widespread, it is difficult to anticipate its consequences. At such early stages, the direction of development of new technologies is usually not fixed and could be influenced relatively easily. However, there is a lack of reliable information on the future development dynamics and impacts of the technology in order to influence it in a targeted manner or, in more modern terms, to ensure its responsible governance (Genus and Stirling 2018). In later phases, when the technology has become widespread and established, control and change are difficult due to lock-ins and path dependencies. The advantages and disadvantages of the technology become clearer as it develops and spreads. Experience is gained, effects can be observed, perhaps even measured. At the same time, the technology becomes socially embedded; it evolves along with, for example, regulations, sunk costs, skills or habits – it becomes institutionalised. This leads to the second part of the dilemma, the control problem. As (Collingridge 1980, 17) puts it:

> The second horn is that by the time a technology is sufficiently well developed and diffused for its unwanted social consequences to become apparent, it is no longer easily controlled. Control may still be possible but has become very difficult, expensive and slow.

For Collingridge, overcoming the first part of the dilemma will require forecasting methods good enough and reliable enough to be used as a basis for more than just warning, but for taking steps to steer the new technology in a desired direction. However, he doubts that such methods will ever be available. He therefore suggests tackling the second horn of the dilemma by avoiding over-dependence on a new technology. This is to maintain a certain degree of flexibility and, therefore, correctability. Progress should be made in smaller, incremental and relatively easily reversible steps, accompanied by constant monitoring to detect negative consequences as early as possible.

The current situation in the field of automated vehicles certainly shows characteristics that correspond to the dilemma outlined by Collingridge (Mladenović et al. 2020). This applies in particular to the information problem. Most experts agree that AVs have the
potential to lead to far-reaching changes in the mobility system (Yen et al. 2024; Cohen and Cavoli 2019). However, it is still an open question whether these developments will lead to a more sustainable transport system (Thomopoulos and Givoni 2015). Both ‘plausible utopias’ and ‘plausible dystopias’ are envisaged (Lyons 2022, 9). AVs can make private car ownership more attractive, thereby reinforcing car-centric mobility patterns, or they can increase the attractiveness of public transport if they are mainly used as robo-taxis. Once the development has moved in a certain direction, self-reinforcing dynamics and path dependencies can become entrenched (Legêne et al. 2020; Schippl, Truffer, and Fleischer 2022; Stilgoe and Mladenović 2022). For example, individually used automated vehicles may become so attractive that they draw passengers away from traditional public transport, weakening its economic viability and modal share, which in turn may lead to reductions in traditional public transport services, further weakening the attractiveness of public transport (‘vicious cycle’ of public transport; (Fraedrich et al. 2017)). Used as highly comfortable single-occupancy vehicles, AVs could thus amplify the negative effects that we have known from cars for decades (e.g. congestion, emissions, urban sprawl). If such path-dependency is established, AVs will make the mobility system less, not more, sustainable (Thomopoulos and Givoni 2015). Status quo car dependency would only be replaced by high-tech car dependency, as described by (Norton 2021).

On the other hand, many experts see automated driving as a central building block of a far-reaching mobility transition, comparable to the energy transition in Germany in terms of its paradigmatic reorientation of a sociotechnical system (Canzler, Knie, and Ruhrort 2019; UITP 2017). In this view, it is considered possible that AVs will contribute to the flexibilisation and expansion of public transport and thus increase the competitiveness of the public transport system as a whole (‘virtuous cycle’ of public transport; (Fraedrich et al. 2017)), especially in rural areas (Schippl and Truffer 2020). By eliminating the driver, costs can be reduced, e.g. by using smaller, fixed-route or more flexible buses, or smaller shuttles or robo-taxis that run at higher frequencies, not replacing traditional public transport but significantly complementing it and thus strengthening the overall system (UITP 2017). The attractiveness and use of the private car would decline and the sustainability of the mobility sector would increase considerably (Legêne et al. 2020).

**Change and transitions of sociotechnical systems**

The lack of information outlined in the first horn or problem of the Collingrid-dilemma is clearly present in the case of AVs. According to current assessments, the further development of AVs is not clearly foreseeable, nor are the future consequences of these developments for mobility and society. The most obvious reason for this is that statements about future developments are always associated with uncertainty. However, this paper will develop the argument that ignorance with respect to current system conditions and processes also plays an essential role in this context. Innovations in the mobility sector encounter a selection environment that is to a large extent already established and that reflects, at least to some extent, the preferences and norms prevailing in society (Granstrand and Holgersson 2020; Rip and Kemp 1998). The relevance of social perceptions, preferences and norms for developments in the mobility sector has
been repeatedly emphasised by authors from social science-oriented mobility research (Cohen et al. 2020; Dennis and Urry 2009; Geels et al. 2012; Milakis and Müller 2021; Sheller 2012).

Not only the innovations themselves, but also many non-technical factors and dynamics that make up the mobility system are decisive for setting the course of development trajectories or path dependencies of innovations. As many authors have shown (Geels et al. 2012; Fraedrich, Beiker, and Lenz 2015), approaches and findings from transition research can help to make such complex sets of interrelationships more tangible and empirically accessible. Transition research has repeatedly shown that innovation processes and transformative change depend on very different factors that go far beyond the technical and economic performance of innovations. The co-evolutionary interplay between technical and societal or institutional factors is of particular interest in transition research (Geels 2004; Markard, Raven, and Truffer 2012). Infrastructures such as mobility or energy are conceived as sociotechnical systems in which elements of different nature are relevant and align to a relatively stable ‘configuration that work’ (Rip and Kemp 1998). The regime is usually understood as the institutional core that forms the deep structure of sociotechnical systems. In this sense, transitions involve changes in these deep structures. Often this understanding of institutions as relatively stable, elastic fibres that stabilise the centre of sociotechnical systems is further specified by drawing on Scott’s understanding of institutions as consisting of regulative, normative and cognitive elements (Scott 2008):

- Regulative institutions refers to formal rules, laws, directive, defined standards;
- Normative institutions refers to values, norms or expectations of how actors are supposed to behave;
- Cognitive institutions refer to shared frames through which meaning is made, common belief and understanding of what is ‘normal’, shared logics of action, taken-for-grantedness.

According to (Hoffman 2001), cited in (Scott 2008), these three pillars can be interpreted as a kind of continuum that extends from the conscious to the unconscious, from the legally enforced to the self-evident. In this sense, and following (Geels 2011, 31), the regime refers to the barely tangible or even unmeasurable cognitive-normative institutions (e.g. engineering beliefs, heuristics, routines, policy paradigms, social expectations and norms), and thus describes a subset of the sociotechnical system. The latter then also includes the more easily tangible or visible elements of infrastructures (e.g. artefacts and infrastructure, regulations, public opinion).

Although regimes are usually characterised by strongly entrenched institutional settings, it has been repeatedly shown that these institutional configurations do not necessarily have to be coherent. Rather, different, perhaps even conflicting rationalities or institutional logics may be sedimented in the regime and influence the design, pace, and direction of innovation (Fuenfschilling and Truffer 2014; Yap and Truffer 2019). For example, in sociotechnical systems, there is usually a conflict between a service-oriented rationality and an economic rationality, as is often shown with hospitals: here, a logic of optimal medical care competes with a logic of economic efficiency. It is obvious that the mobility sector is particularly exposed to incoherent institutional
logics (Cohen 2010; Drexler et al. 2022). For example, the different modes of transport are repeatedly competing for resources such as money, public support or road space (Schöller-Schwedes 2010). The old conflict between actors who frame car traffic as environmentally harmful and worth avoiding and those who do not like to conceive of mobility without their own car and the corresponding infrastructures is also well known (Canzler et al. 2018; Dennis and Urry 2009).

So, it can be learned from transition research that the potentially heterogeneous set of cognitive and normative institutions is interrelated with the development trajectories of new technologies. To anticipate the future course of AVs development, it is necessary to understand the normative-cognitive institutional configurations that co-determine the directionality of the relevant regime. The concept of directionality was introduced by (Stirling 2008; 2009) as part of his call to open up the process of technical change to alternative options. It is based on the idea that sociotechnical change has a direction, that decisions are made between different directions and that actors in the innovation system gradually become ‘blind’ to alternatives (Schot and Kanger 2018; Weber and Rohracher 2012). Based on this understanding and following Yap and Truffer (2019) and Schippl and Truffer (2020), the concept of directionality is understood here as the alignment of the system configuration that co-determines the transformative potential of an innovation. Directionality describes the state of the regime in relation to an innovation, i.e. the degree of openness/closedness to certain development directions and path dependencies. Weber and Rohracher (2012) describe ‘directionality failures’ as a recurring cause of undesirable developments in sociotechnical transitions. This is the case when an innovation that seems promising from a sustainability perspective ultimately only leads to the stabilisation of unsustainable developments or conditions. Weber and Rohracher (2012) identify the development of ‘shared visions’ as an important approach to counteracting such undesirable developments. Such normative visions then provide the necessary orientation for a coherent policy portfolio to emerge. The authors also address the problem that consensus on the right development direction in the sense of a ‘shared vision’ is often not readily available, but must first be developed through appropriate policies and processes.

Against this background, it becomes clear that the governance of innovation and transition processes requires knowledge that is approached in different epistemic ways. This includes knowledge about the (future) consequences of an innovation, which is primarily addressed by the first Collingridge challenge. Obviously, good knowledge of important conditions and causal chains in the current system is required. These are only partly directly accessible empirically. For example, research on the configurations and dynamics of cognitive-normative institutions requires more interpretive, hermeneutic approaches. In addition, knowledge is needed about the prevailing target systems (e.g. phase-out of internal combustion engines by 2035) and the underlying patterns of reasoning.

**Classification of TA-knowledge in context of sociotechnical transitions**

For some time now, TA and transition research have been discussing that different types of knowledge are required in order to understand transitions and to be able to successfully steer or at least influence them. In the beginning, a distinction was made between
three types of knowledge (Grunwald 2020; Hirsch Hadorn 2005; WBGU 2011; Weber and Whitelegg 2003)

- Systems knowledge (also called explanatory knowledge);
- Orientation knowledge (also called target knowledge);
- Knowledge for action (also called transformative knowledge).

The category of orientation knowledge is particularly problematic because it combines very different epistemological approaches and bodies of knowledge (Grunwald 2020). On the one hand, this category includes knowledge about the future, i.e. knowledge about possible future developments that is largely based on if–then assumptions. Such knowledge is generated or expressed, for example, through explorative scenarios or Delphi methods, which are often based on an aggregation of expert assessments. On the other hand, the category of orientation knowledge also includes normative knowledge, i.e. knowledge about goals for certain sectors/developments and their rationale (Hirsch Hadorn 2005). Such goals or goal systems are based on societal value systems and are derived from legal and/or ethical reflections (Grunwald 2020). Prominent examples are goals related to climate protection, sustainability concepts with their indicators or equity issues as expressed in concepts such as energy justice or fair access to mobility options. Even if both types of knowledge can be combined methodologically, for example in normative scenarios or by differentiating between plausible and desirable developments, they are still epistemologically distinct bodies of knowledge that can have very different meanings in transition processes. Against this background, Grunwald and Schippl (2013) propose a division into four types of knowledge, which then distinguish between normative knowledge, system knowledge, future knowledge and knowledge for action.

Some years ago, Grunwald identified a fifth category of knowledge, which he calls hermeneutic knowledge (Grunwald 2020; 2014). This type of knowledge aims to understand the meaning that innovations and related images of the future and processes of change have for particular individuals, groups or societies. The aim is to examine textual, graphic or filmic statements or representations for the motivations behind them. In a very similar way, approaches from transition research aim to illuminate the conglomerate of intangible cognitive and normative institutions that decisively mould the respective regime configuration (Geels 2004; Fuenschilling and Truffer 2014) and its directionality (Schippl and Truffer 2020).

In this paper, we propose a slight shift or specification of these categories, as described in (Grunwald 2020, 112), which is less concerned with their content than with the relationship between the categories (Table 1). The core argument is that ultimately all knowledge that seeks to inform transition processes can be

<table>
<thead>
<tr>
<th>Knowledge categories</th>
<th>Motivation / main objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical systems knowledge</td>
<td>to understand, to explain</td>
</tr>
<tr>
<td>Prospective knowledge</td>
<td></td>
</tr>
<tr>
<td>Hermeneutic knowledge</td>
<td></td>
</tr>
<tr>
<td>Normative knowledge</td>
<td>to guide, to orient</td>
</tr>
<tr>
<td>Instrumental knowledge</td>
<td>to design, to implement</td>
</tr>
</tbody>
</table>
described as orientation knowledge in a broader sense. Only the category of normative knowledge is geared towards ‘direct’ orientation, in the sense of determining desired or undesired directions of development with more or less specific target areas. It thus differs from the types of knowledge that are primarily oriented towards understanding (or explaining). These include the three categories of empirical system knowledge, hermeneutic knowledge and also future knowledge, which, although they describe different ways of knowing, are ultimately directed towards better understanding the systems interrelations and therefore also further possible developments (Table 1). Even if these types of knowledge with their corresponding ways of knowing can never be free of normative settings, the primary goal is a better understanding of system-relevant interrelations or conditions. This kind of knowledge is at the heart of the so-called information problem in the Collingridge-dilemma: impacts cannot be easily foreseen as long as the technology is not yet sufficiently developed and widespread.

**Knowledge on the transformative potentials of AVs**

In the following, the state of knowledge in the different knowledge categories for the situation in the field of AVs is roughly outlined. The aim is not to provide a comprehensive overview of the literature in each knowledge category. Rather, the knowledge types are described in terms of the knowledge that is essential for the governance of AVs and specific challenges are identified. The focus is on the categories of prospective knowledge, normative knowledge and hermeneutic knowledge, as these are particularly relevant to the argument of this paper.

Table 2 shows which elements of knowledge within the five categories are particularly relevant for understanding and anticipating the transformative potentials of AVs and for enabling governance with regard to a socially desired path dependency.

<table>
<thead>
<tr>
<th>Knowledge category</th>
<th>Relevant aspects for responsible governance of AVs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical systems knowledge</td>
<td>• Understanding cause-effects in the socio-technical system of mobility</td>
</tr>
<tr>
<td></td>
<td>• Empirical evidence on mobility behaviour</td>
</tr>
<tr>
<td></td>
<td>• Empirical evidence on the acceptance of AVs</td>
</tr>
<tr>
<td>Prospective knowledge</td>
<td>• Estimation (modelling) of future traffic and social effects (of AVs)</td>
</tr>
<tr>
<td></td>
<td>• Understanding the scope of plausible socio-technical developments, scenarios (in relation to AVs)</td>
</tr>
<tr>
<td></td>
<td>• Assumptions on future acceptance of measures/technologies/services</td>
</tr>
<tr>
<td>Normative knowledge</td>
<td>• Knowledge on the relevant normative configurations in the regime</td>
</tr>
<tr>
<td></td>
<td>• Knowledge on visions, goals for mobility/AVs and on its reasoning</td>
</tr>
<tr>
<td></td>
<td>• Criteria/indicators or arguments for assessment (e.g. sustainability indicators, capability-approach, ethical reflection on conflicting goals)</td>
</tr>
<tr>
<td>Hermeneutic knowledge</td>
<td>• Understanding relevant cognitive-normative institutions with their backgrounds, their history, their deeper social embedding</td>
</tr>
<tr>
<td></td>
<td>• Understanding fundamental rationalities and/or the deeper embedding of ‘institutional logics’ which co-shape directionality</td>
</tr>
<tr>
<td></td>
<td>• Understanding the deeper settings which influence social legitimacy of changes (caused by AVs)</td>
</tr>
<tr>
<td>Instrumental knowledge</td>
<td>• Identification of needs for action</td>
</tr>
<tr>
<td></td>
<td>• Knowledge of effectiveness, efficiency and acceptance of measures</td>
</tr>
<tr>
<td></td>
<td>• Process knowledge on how to carry out (participative) planning</td>
</tr>
</tbody>
</table>
**Empirical systems knowledge**

In terms of methodology, this is about ways of gaining knowledge that are based on observations, surveys and modelling. The objective is a good understanding of the current mobility system. The various surveys and user questionnaires that are regularly conducted in Germany (Nobis and Kuhnimhof 2018) contribute to this. Transport modelling is a well-established tool for simulating traffic flows and the potential impact of interventions in the mobility system at different spatial and administrative scales. Various other disciplines contribute to a better understanding of cause–effect relationships in the mobility system. In addition to engineering and economics, these include the increasingly important political and social sciences (Schöller-Schwedes 2010; Ryghaug et al. 2023).

Obviously, a good knowledge of the technological side of AVs with its innovation system or ecosystem (Granstrand and Holgersson 2020) is required, which also extends to aspects of regulation and policy options (Yen et al. 2024). Likewise, research is needed on the various dimensions of societal and social acceptance of AVs, as it has been addressed for several years in sometimes quite differently designed studies (Gkartzonikas and Gkritza 2019; Duboz et al. 2022). Not all studies in this area come to the same conclusions, which is certainly due to different methodological approaches and different framing of the object of investigation. However, many of these studies indicate that high expectations of the usefulness and quality of automated services can at least partially outweigh existing safety concerns (dos Santos et al. 2022). In general, younger males with higher levels of education are more open to purchasing/using vehicles with automated driving functions (CEC 2020).

**Prospective knowledge**

Prospective knowledge is based on assumptions about developments and cause–effect relationships that lie in the future and are therefore not directly observable. It helps to understand the space of possible and plausible future developments and their potential consequences or impacts. As such, prospective knowledge is closely linked to systems knowledge. However, foresight experts repeatedly point out that knowledge about the future should not be based too heavily on findings, developments or connections from the past (Miller 2007). It is important to consider less likely or disruptive developments as well. Furthermore, it is repeatedly emphasised that anticipating future developments, producing forecasts or scenarios, is not an end in itself. Rather, such knowledge of the future should contribute to better decisions in the present because it helps to better understand the possible or plausible consequences of decisions (Miles and Keenan 2003; Urry 2016). This type of knowledge and the methods to generate it are well established in mobility planning and policy. Typical examples are forecasts of future transport demand or estimates of the impact of new infrastructure such as roads or cycle lanes. Transport modelling often plays a key role here. Various forms of citizen and stakeholder participation can play a role both in planning processes and in the evaluation of potential future impacts (Rosa et al. 2021).

As regards AVs, the main task is the anticipation of plausible future development directions of AVs and their potential interrelations with and impacts on mobility and society. In particular, the potential impact of different forms of AVs on mobility
systems and spatial development has been widely studied. With regard to passenger transport, a large number of often model-based studies have already addressed the question of whether or to what extent AVs will affect transport demand and mileage (BMVI 2017; Meyer et al. 2017; WEF and BCG 2018). Many studies assume an increase in transport demand and mileage when automation levels 4 and 5 are reached (Bösch et al. 2018; Friedrich, Hartl, and Magg 2018). Overall, the available research makes it very clear that the impact on transport depends very much, on whether AVs are used individually or collectively (see section 2.1).

In general, the studies mentioned above assume that one or the other form of AVs has been commercialised to some extent. On this basis, the traffic models calculate impacts in an if–then logic. How these different forms of AVs might come to exist is rarely explored from a sociotechnical perspective. Few studies explicitly seek to understand what factors influence how more or less sustainable co-evolutionary dynamics might emerge and be transformed into path dependencies. However, there are some examples: Based on focus groups and interviews, Galich and Stark (2021) find that whether AVs become more or less sustainable depends very much on factors such as pricing structures, regulatory requirements and the design of AVs services. Fraedrich, Beiker, and Lenz (2015) use the multi-level-perspective to assess how different transition pathways might unfold and shape the mobility sector. Schippl, Truffer, and Fleischer (2022) explicitly aim to better understand situations in which alternative pathways of transformative change are likely to emerge. They focus on potential variations of different institutional logics in the mobility sector to analyse how these variations might co-evolve with AVs and induce more or less sustainable path dependencies and lock-ins. The approach to shed more light on normative-cognitive institutions moves in the direction of hermeneutic knowledge (see below), but is listed here with prospective knowledge because of the overarching goal of deriving if–then scenarios. Also called for by some authors and important for a sustainability assessment is a more far-reaching consideration of the consequences of AVs beyond the mobility sector (Cohen et al. 2020). As illustrated in (Mladenović 2019), impacts of AVs may well go beyond the mobility sector, extending to aspects such as everyday activities or social values and norms.

It is certainly in the nature of the matter that prospective knowledge in the field of AVs is subject to many uncertainties. It is open which form of AVs will dominate future mobility. A proliferation of robo-taxis and shuttles could lead to significantly more sustainability, while a massive dominance of private AVs could even weaken sustainability. It has been shown several times that plausible scenarios can be developed for both development paths (Legêne et al. 2020). Nevertheless, to a certain extent, AVs will have to adapt to existing infrastructures and established mobility needs. Fraedrich, Beiker, and Lenz (2015) emphasise that the emergence of new technologies, such as AVs, is embedded in an existing socio-technical system, which requires a certain alignment of the new technology with the established environment. The range of future options in the area of AVs and the plausible scope of a corresponding scenario funnel can therefore be delimited with some certainty.

**Normative knowledge**

It is essential to use normative criteria or targets to distinguish between desirable and undesirable developments when assessing the future trajectories of AVs and their
various impacts. From a research perspective, this often includes the development or normative justification of evaluation criteria, which can be based on concepts such as sustainability and/or ethical reflections (Roache 2008), the latter especially when it comes to weighing up conflicting goals or trade-offs. The fundamentally important question of the extent to which academia should play a more active role in the development and implementation of visions, goals and policies is beyond the scope of this article.

It is important to note that normative criteria or assessments do not only relate to the properties of new technologies and processes, but rather to the way in which these artefacts and processes are used or expected to be used in the respective sociotechnical system (Bijker and Law 1992; Grunwald 2022). It is not the technology itself that is sustainable or not, but the way it is used and embedded in the sociotechnical system. What the assessments are really about are the interrelationships of AVs with and/or impacts on the sociotechnical system of mobility. Relevant questions are whether safety will improve or not, emissions will be reduced or not, passenger kilometres travelled will increase or not or whether modal choice will be changed or not. None of these impact categories are in principle new, since mobility is a long-established sociotechnical system. Therefore, normative assessments of the development of AVs are inevitably linked to ideas about what desirable mobility options and a desirable mobility future should look like.

When it comes to what exactly are desirable developments in the mobility system, opinions often diverge significantly. In Germany, as in many other countries, it can be observed that at a very general level, goals such as sustainable and/or climate-friendly mobility are widely accepted. How exactly these goals are to be implemented, or which sub-targets or milestones are to be set, however, often remains controversial (Drexler et al. 2022; Lambrecht et al. 2023). Mobility is considered a central prerequisite for quality of life and economic development in modern societies. At the same time, every form of motorised mobility is also associated with negative consequences for society, such as pollutant emissions, energy consumption, land use or the risk of accidents. In this field of conflicting goals, normative orientation to guide policy is enormously important, but not easy to achieve (Schöller-Schwedes 2010). As mentioned in Section 2, the regime in the mobility sector is characterised by at least two competing institutional logics: according to one logic, driving is considered indispensable, while according to the other logic, car use should be strongly reduced, especially in cities. This is illustrated by the conflicts over how to deal with car traffic, some of which have been going on for decades (Canzler and Knie 2016; Mögele and Rau 2020). Typical issues are: speed limits on motorways, reversing traffic lanes in favour of bicycles in cities, car-free zones in cities, or removing parking spaces in residential areas to improve the quality of life (Kirschner and Lanzendorf 2020).

Many observers point to such challenges for the development of a shared vision or target system for the German mobility sector (Schipp and Arnold 2020). The German Association of Cities (Horn, Kiel, and von Lojewski 2018) sees the development of a shared vision of future mobility as a necessary precondition for a mobility transition. Haas (2020) comes to the conclusion that windows of opportunity for less car-orientated mobility do exist, but that they do not have enough political backing for achieving far-reaching impacts. Drexler et al. (2022, 1) find in their research on how incumbent actors of automobility in Germany have framed the issue of a transition of mobility and transport:
This paper demonstrates that there is no common understanding of the problems and solutions to foster a mobility transition, as the diversity of problems and solutions proposed within the frame elements is high and complicates the prevailing implementation gap of the mobility transition.

There are targets for decarbonisation and for the integration of low-carbon fuels in Germany (an in the EU). The existing Climate Protection Act sets long-term targets for the transport sector. The transport sector must reduce its GHG emissions by 48 percent by 2030 compared to 1990. However, its effectiveness is weak because the Climate Protection Act cannot achieve the necessary transport planning management (Horn 2021). A related goal is to have 15 million electric cars in the vehicle fleet by 2030. However, this is a fuel and powertrain transition and not a broader transport or mobility transition (Manderscheid 2020), which includes changes in mobility behaviour. Horn (2021) notes that, unlike energy legislation, transport legislation lacks targets, as well as horizontal, cross-modal planning and vertical coordination between different levels of government. This statement is supported by a 2018 OECD report, which found that Germany’s transport sector lacks an overarching policy strategy (OECD 2018). And also the OECD’s report from 2023 calls for a more pro-active transport planning in Germany, that moves from a kind of ‘predict and provide’ approach towards strategic planning (OECD 2023). The Federal Transport Infrastructure Plan (BVWP) is the most important instrument of federal transport infrastructure planning in Germany. This approach is often criticised because it predicts demand and provides the corresponding infrastructure rather than pursuing supply-oriented transport planning (Siebert 2022).

A clear and consistent orientation of the mobility sector towards reducing car ownership and use has therefore not yet been comprehensively established. Against this background, it appears difficult to channel the future development paths of AVs in a specific direction. The regulatory framework for AVs therefore remains vague. Certainly, recently enacted legal regulations in Germany support the use of AVs in the public transport sector (Deutscher Bundestag 2021). The regulation requires L4 – vehicles to be monitored by a technical operator who can intervene remotely if necessary. An operator can monitor several vehicles at the same time, which means that the law is tailored to fleet operators as they are established in the public transport sector. However, the law is not linked to clear targets for the introduction of AVs or even integrated into an overarching strategy for the mobility of the future. In addition, in the private car sector, automation is also on the way and is being commercialised to a certain extent.

**Hermeneutic knowledge**

Hermeneutically derived knowledge helps to understand the motivations, values and reasoning patterns underlying the above-mentioned conflict constellations and the corresponding normative ambiguity. This type of knowledge aims to identify and understand widespread or at least influential cultural beliefs or expectations with their historical background, the motivations and interests of actors, or also the background of acceptance constellations, and thus make them usable for orientation in transition processes (Grunwald 2014; 2020).

Hermeneutic knowledge taps into cognitive and normative institutions that shape regime configurations and their stability/changeability, which are difficult to access...
with ‘classical’ empirical methods such as surveys. The exploration of knowledge in this field must therefore be highly interpretative. The aim is to understand fundamental rationalities or institutional logics (Fuenfschilling and Truffer 2014) in the regime, which in their interplay co-determine the directionality of the sociotechnical system in relation to AVs (Schippl and Truffer 2020). Mobility research is often dominated by quantitative approaches and there is a tendency to underestimate how relevant such hermeneutically accessible aspects can be for developments in the mobility sector (Mögele and Rau 2020). Yet the design of the mobility regime is co-determined by many implicit ideas/expectations of what is right or appropriate. It is important to investigate the deeper social embedding of these ideas and expectations to better understand the conditions for relieving the aforementioned lack in normative clarity. Ruhrort (2023) emphasises the importance of narratives in sustainable transitions and that competing narratives can be observed in Germany, struggling what a mobility transitions should look like. In the section on prospective knowledge, we referred to a study by Schippl, Truffer, and Fleischer (2022), which shows that the institutional dynamics in the current mobility sector deserve more attention, as they may be the starting point for quite different path dependencies. It is therefore important not only to look at possible future consequences of AVs, but also to understand the directionality of the current regime. Hermeneutic knowledge, with its focus on the institutional configuration that exists today, can help to better understand these initial conditions for possible path developments.

In particular, when it comes to the meaning of the car, there are studies that can certainly be assigned to the field of hermeneutic knowledge. For example, in his book ‘The Psychology of the Car’, Gössling (2017) works out that many of our social norms and perceptions are influenced by popular culture, including films, literature, music and games, and that perceptions and attributions of meaning to the car are also shaped in this way. Mögele and Rau (2020) use a discourse analytical approach to examine the extent to which the dominant mobility culture in two German states strongly influenced by the automotive industry is shaped by the identity-forming image of an automotive state. The authors are able to show that this image inhibits change towards a more sustainable, less car-centred mobility culture. To cite another example: Groth et al. (2023) investigated the extent to which everyday mobility is portrayed in the media and the extent to which aspects of a mobility transition are addressed. The study focused on various entertainment programmes, including a popular crime series. The authors found that the current controversy about the future role of the car is also reflected in the media, with car-centred storylines dominating. In a similar vein, Martin (2021) examines AVs visualisations, not to anticipate future developments, but to better understand the discursive meaning carried by these visualisations.

**Instrumental knowledge**

This category is about the efficiency and implementability of measures (Givoni 2014). Thus, this knowledge type builds strongly on the other categories. It is here that the lack of normative orientation becomes particularly relevant. Efficiency and effectiveness of measures can only be assessed with regard to specific targets. Normative and hermeneutic aspects are also decisive in issues of acceptance, which always involve an implicit
or explicit weighing of perceived advantages and disadvantages of a technology or a new service. Further, this type includes process knowledge and practical knowledge as required for the planning and implementation of specific measures.

**AVs as a case of Mode-2-directionality**

In this section, the five types of knowledge are used to distinguish the specific situation of AVs from other areas of innovation. The intention is to highlight that AVs represent a specific mode of directionality that requires a specific mode of governance. Central characteristics of AVs are, first, the confined scenario funnel in the realm of prospective knowledge and the uncertainty about what kind of co-evolutionary dynamics will prevail within the funnel. Second, the situation is characterised by a lack of normative clarity and the (related) need for more hermeneutic knowledge. Directionality is confined, but is characterised by heterogeneous institutional logics within the domain bounded by the funnel’s boundaries.

Figure 1 summarises this characterisation of AVs and contrasts it with two other modes of directionality that can be observed in the mobility sector, here represented by electric mobility and by blockchain technology. It is inspired by a categorisation introduced by Armin Grunwald, concerned with the fact that future-oriented studies often come to different conclusions, which raises the question of how TA can provide orientation for political decisions on the basis of such heterogeneity (Grunwald 2014). Depending on the extent to which future studies converge or diverge in a given field, Grunwald distinguishes between three modes of orientation. Mode-1-orientation is characterised by a fairly good knowledge of the relevant conditions and cause–effect relationships in the respective sociotechnical field. Use cases for the new technology are quite well understood and can be narrowly defined. Experts’ assessments of the

![Figure 1](image-url)  
**Figure 1.** Different modes of directionality.
future development of the innovation tend to converge. Mode-2-orientation is characterised by moderate or bounded diversity. There is clear uncertainty about the development directions of the innovation. However, the potential use cases and development paths remain within a reasonably definable range of possibilities. Mode-3-orientation is confronted with ‘unbounded divergence’ (Grunwald 2014, 9). There is a high degree of uncertainty and normative diversity regarding the possible applications and impacts of the innovation. Assessments of possible future developments and their impacts diverge significantly.

This categorisation can be extended to distinguish the three different modes of directionality shown in Figure 1. Grunwald’s modes take the state of knowledge or uncertainty as the key criteria for assigning technologies to the modes. Accordingly, the situation in the field of prospective knowledge is the point of reference. This is represented by the x-axis in Table 1. For a more detailed representation of the configuration of knowledge in a mode, we propose to explicitly add another dimension which is related to normative and hermeneutic knowledge: the homogeneity/heterogeneity of the relevant institutional configuration, which is reflected on the x-axes. It should be noted that the positioning on the x-axis and the y-axis in Table 1 influence each other. E-mobility can be taken as an example for Mode-1-directionality. The technology’s scope of application is rather narrow. The question is how quickly e-cars will become widespread and not how they will be deployed and used. It is obvious that they will be used in the same way as cars are used today. At the same time, in Germany (and in the EU), e-mobility is largely seen as a desirable development in the political arena and is supported by corresponding regulations. There are concrete targets for market penetration and for the expansion of the charging infrastructure, which are shared and supported programmatically by a broad range of relevant actors. This must of course be taken into account in the area of prospective knowledge when forecasting future market penetration. The further diffusion of e-mobility is likely not only because it is technically and economically feasible, but also because it is socially desirable. The technology falls into Mode 1, where systems knowledge and instrumental knowledge are particularly important to enable the politically favoured rapid diffusion of electric vehicles.

If we had the same level of normative homogeneity in the area of AVs as we have it for e-mobility, AVs would have to be moved significantly more into Mode 1 in Figure 1. For example, a future in which AVs are banned in private transport and only allowed in public transport (with technical supervision, as currently provided for by legislation) would significantly narrow the scope for future development paths. The scenario funnel would become much smaller. Due to the heterogeneous regime configuration, however, it is currently questionable whether, or under what conditions exactly, such a strong political support for a certain direction could develop. We therefore need more research and knowledge in the areas of normative and hermeneutic knowledge in such a situation of Mode-2-directionality.

At the same time, AVs are clearly distinguished from technologies from the area of Mode-3-directionality, whose potential applications are very open and hardly confinable. Typical examples are nanotechnology or artificial intelligence. Blockchain technology should be cited here as a current example of Mode 3 in mobility. The hype around blockchain technology, which was mainly triggered by the enormous increase in the price of bitcoin and other cryptocurrencies, has led to the development of a large number of ideas
and projects in various areas, including mobility. Some of these ideas overlap with current trends such as mobility-on-demand, autonomous driving, and vehicle electrification (Gösele and Sandner 2019; Karger, Jagals, and Ahlemann 2021). In interaction with these trends, blockchain could have a disruptive effect on the sociotechnical system of mobility in various ways. However, the possible directions of development and their impacts are not yet well understood and difficult to narrow down. In contrast to AVs, i.e. Mode 2, no reasonably delimitable scenario funnel can be identified. On the side of technology development, there are still hardly any clear reference points towards which the range of future options could be determined or narrowed down. In Mode 3, the main concern must be to improve the basis for timely democratic deliberation in the respective field of action (Grunwald 2014). More reliable knowledge about potential future applications is needed as a basis for normative orientation. Hermeneutic knowledge can help to better understand the causes and sources of divergent ideas and assessments of potential future applications.

Discussion and conclusions

As this paper has shown, the major challenge for the governance of AVs does not lie in the dilemma outlined by Collingridge or in the corresponding area of prospective knowledge. The Collingridge-dilemma is strongly informed by a consequentialist paradigm that sees uncertainty as a central challenge to the responsible governance of new technologies with potentially far-reaching societal impacts. The development of a complex technology such as AVs certainly involves many uncertainties, not only in terms of its technical potential, but especially in terms of its far-reaching implications for mobility and society. Nevertheless, research to date provides a fairly good understanding of the key influencing factors and possible cause-and-effect chains associated with AVs developments, at least as far as possible impacts on transport are concerned. There are good reasons to believe that only the collective use of AVs will effectively support a sustainable transition of the mobility system. At the same time, it is reasonable to assume that AVs can facilitate attractive collective mobility services, especially in less densely populated spatial categories. Equally clear is the risk that increased individual use of automated vehicles will exacerbate the negative effects of the mobility system as a whole – probably with the exception of safety issues.

Uncertainty about the consequences of AVs is certainly a challenge. More research and knowledge in this area is certainly necessary and useful. However, even if there were perfect knowledge of the consequences, another, probably even more serious problem would remain. For AVs, and more generally for mobility, it is not entirely clear what is a desirable development or consequence and what is not. Responsible governance of an ambivalent Mode-2 technology such as AVs inevitably requires more normative and hermeneutic knowledge. The main problem is the lack of a binding vision for the mobility of the future, supported by a clear system of objectives towards which the development of AVs could be directed. What is missing is a broad societal consensus on sustainable mobility, as called for, for example, by the German Association of Cities (Horn, Kiel, and von Lojewski 2018). In this normative vacuum, there is a risk that AVs will simply stabilise the existing structures of the mobility system or even increase the attractiveness of the car at the expense of public transport. If the latter is
to be countered by policy measures, socially legitimised objectives are needed. This also means recognising more explicitly that the development of AVs should not and cannot be depoliticised. Or, as Mladenović et al. (2020, 13) put it, ‘that technological development is irreducibly a political and value-driven choice rather than an instrumental facilitation of an inevitable (automated) future’. In this context, Stilgoe and Mladenović (2022) point to Winner’s concern that the relevance of politics is often recognised too late in the course of technological progress, and that we sleepwalk through technological revolutions (Winner 1980).

Too much focus on prospective knowledge, in the sense of the first part of the Collingridge-dilemma, tends to distract from the main problem. The necessary broader political and societal commitment to a sustainable mobility future should be created proactively, not reactively. It may be considered irresponsible to postpone the development of a widely shared vision into the distant future, when it may be more difficult or even too late to implement it. Norton emphasises that AVs are tools that equip us to solve our problems ourselves (Norton 2021). In this sense, the core message of this paper can be summarised with a quote from Norton’s book ‘Autorama’: ‘Let’s ask what future we want and need, and then talk about the technology we need to get there’ (Norton 2021, 225). Ideally, directionality is already configured today in such a way that unsustainable path dependencies of AVs become unlikely – to avoid directionality failures in the sense of Weber and Rohracher (2012). Similarly, Haugland and Skjølsvold (2020) argue that policy makers should be more active in setting desirable directions for transport automation. This requires normative orientation that is robustly accepted by society. The necessary social discourses and ethical considerations can or should be taking place today (Roache 2008). Normative knowledge requires not only visions and goals, but also processes that allow for the socially robust development of such visions and goals. In TA and related fields, there is a wide range of experience with participatory approaches to bring together different stakeholders and to identify potentials and obstacles for the development of common visions in a given field (Rosa et al. 2021). At the very least, promising approaches should be available here to better address a normative deficit in mobility.

It can be concluded that the application of the five types of knowledge in the case of AVs is helpful to elaborate which knowledge deficits should be brought to the fore in order to make AVs usable for a sustainable transition of the mobility sector. The approach helps to point out the lack of normative orientation and to sensitise society and politics for this deficit. It also justifies a stronger focus on hermeneutic approaches in research, which are often reduced to accompanying research and thus remain on the fringes rather than at the centre of research projects. TA-related research on AVs needs to focus more explicitly on the relationship between the new technology and the established regime with its values and expectations. As shown above, TA, at least in a Mode 2 case such as AVs, should not only focus on possible future path dependencies with their various consequences (scenarios, modelling). Rather, it is a matter of better understanding which development paths are possible or probable under today’s initial conditions, and which ideas for a future mobility system are considered desirable – and for what reasons. Here, hermeneutic knowledge can help to better understand the deeper ration-
Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Jens Schippl is Senior Researcher in the “Mobility Futures” research group at the Institute for Technology Assessment and Systems Analysis (ITAS), Karlsruhe Institute of Technology (KIT), and lecturer at the ETH Zurich in the context of the MAS/CAS “Future Mobility”. His research focuses on sociotechnical processes of change, foresight and technology assessment in the mobility sector. He is particularly interested in the development of automated driving and its interactions with path-dependencies in the mobility system.

ORCID

Jens Schippl http://orcid.org/0000-0003-3169-6365

References


Mladenović, Milos, Dominic Stead, Dimitris Milakis, Kate Pangbourne, and Moshe Givoni. 2020. “Sociotechnical Imaginaries of Connected and Automated Vehicle Technology: Comparative Analysis of Governance Cultures in Finland, Germany, and the UK.”


Stirling, Andy. 2009. Direction, Distribution and Diversity! Pluralising Progress in Innovation, Sustainability and Development”.


