

Noortje Marres; Philippe Sormani

## Testing 'AI': Do We Have a Situation? A Conversation 2023

<https://doi.org/10.25969/mediarep/20044>

Veröffentlichungsversion / published version  
Buch / book

### Empfohlene Zitierung / Suggested Citation:

Marres, Noortje; Sormani, Philippe: *Testing 'AI': Do We Have a Situation? A Conversation*. Siegen: Universität Siegen 2023 (Medien der Kooperation 28). DOI: <https://doi.org/10.25969/mediarep/20044>.

### Erstmalig hier erschienen / Initial publication here:

<https://dspace.ub.uni-siegen.de/handle/ubsi/2525>

### Nutzungsbedingungen:

Dieser Text wird unter einer Creative Commons - Namensnennung - Nicht kommerziell - Keine Bearbeitungen 4.0/ Lizenz zur Verfügung gestellt. Nähere Auskünfte zu dieser Lizenz finden Sie hier:

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

### Terms of use:

This document is made available under a creative commons - Attribution - Non Commercial - No Derivatives 4.0/ License. For more information see:

<https://creativecommons.org/licenses/by-nc-nd/4.0/>



MEDIEN DER  
KOOPERATION

Universität  
Siegen

DFG

---

## Testing 'AI': Do We Have a Situation?

A Conversation

Noortje Marres\* & Philippe Sormani<sup>†</sup>

*\*University of Warwick & <sup>†</sup>University of Lausanne*



---

**WORKING PAPER SERIES | NO. 28 | MAY 2023**  
Collaborative Research Center 1187 Media of Cooperation

## **Working Paper Series Collaborative Research Center 1187 Media of Cooperation**

Print-ISSN 2567-2509

Online-ISSN 2567-2517

DOI [doi.org/10.25819/ubsi/10332](https://doi.org/10.25819/ubsi/10332)

[dspace.ub.uni-siegen.de/handle/ubsi/2525](https://dspace.ub.uni-siegen.de/handle/ubsi/2525)

URN [urn:nbn:de:hbz:467-25255](https://nbn-resolving.org/urn:nbn:de:hbz:467-25255)



This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

This Working Paper Series is edited by the Collaborative Research Center Media of Cooperation and serves as a platform to circulate work in progress or preprints in order to encourage the exchange of ideas. Please contact the authors if you have any questions or comments. Copyright remains with the authors.

The Working Papers are accessible online at:

<https://www.mediacoop.uni-siegen.de/de/publikationen/working-papers-media-of-cooperation/>

Print copies can be ordered by sending an email to:

[workingpaperseries@sfb1187.uni-siegen.de](mailto:workingpaperseries@sfb1187.uni-siegen.de)

Publication is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 262513311 – SFB 1187.

Cover image: *Ghost #8 (Memories of a mise en abîme with a bare back in front of an untamable tentacular screen)*, experimenting with OpenAI Dall-E, Maria Guta and Lauren Huret (Iris), 2022.(Courtesy of the artists)

Layout: Anna K. Büdenbender

Universität Siegen

SFB 1187 Medien der Kooperation

Herrengarten 3

57072 Siegen, Germany

<https://www.mediacoop.uni-siegen.de/sfb1187.uni-siegen.de>

[workingpaperseries@sfb1187.uni-siegen.de](mailto:workingpaperseries@sfb1187.uni-siegen.de)

---

## Testing 'AI': Do We Have a Situation? A Conversation<sup>1</sup>

Noortje Marres\* & Philippe Sormani<sup>†</sup>  
*\*University of Warwick & <sup>†</sup>University of Lausanne*

---

**Abstract** This working paper is based on the transcription of a recent conversation between the authors, regarding current instances of the real-world testing of “AI” and the “situations” they have given rise to, or as the case may be, not. The conversation took place online, on the 25th of May 2022, as part of the Lecture Series Testing Infrastructures, organized by the Special Collaborative Research Center (CRC) 1187 “Media of Cooperation” at the University of Siegen, Germany. This working paper is an elaborated version of this conversation and is organised as follows. The introduction presents an expanded version of the lecture abstract, which was used to advertise the online conversation. The bulk of the working paper reproduces and extends the transcribed conversation, as well as parts of the ensuing discussion with the audience. In the conclusion, we address the question “do we have a situation?,” when it comes to AI testing in society, in the light of the conversation, and reflect on “what’s next” in social studies of “AI” testing situations, as well as on turn-taking in (online) conversation.

**Keywords:** AI, social deficit of technology, situatedness, social studies of testing, ethnomethodology, everyday artefacts

### Introduction

Proponents of the “New AI” in computer science as well as in social science and the humanities have claimed that today’s very large deep learning models exhibit radically new capacities for contextual judgement and decision-making as well as situational awareness (see Fig. 1 for a playful reflection on this claim). These arguments are advanced through high profile publications, conference proceedings and arxiv papers (LeCun et al, 2015; Bommasani et al, 2021; for a discussion, see Roberge and Castelle, 2021), but equally through tests and demos, such as DeepMind’s

presentation in the Radcliffe Observatory in Oxford (Hassabis, 2016); AlphaGo’s victory at the Four Seasons Hotel in Seoul, South Korea (Sormani, 2018; Mair et al., 2021) and street trials of self-driving vehicles in urban centers like Phoenix (Arizona) and Coventry in the UK (Marres, 2020). Such public demonstrations have played a notable role not only in the propagation of the claim that the new AI possesses situational intelligence, but equally in the problematization of such claims, as we will discuss here. Below we discuss how social studies of AI testing do, can, and should approach and engage with claims to the new AI’s situational intelligence.

---

<sup>1</sup> We thank Johannes Schick (JS) for convening and Carolin Gerlitz (CG) for chairing the conversation on which this working paper is based, as well as the participants for attending and contributing to it. For further details, see <https://www.mediacoop.uni-siegen.de/de/veranstaltungen/ringvorlesung-testing-infrastructures-noortje-marres-philippe-sormani-testing-ai-a-conversation/>

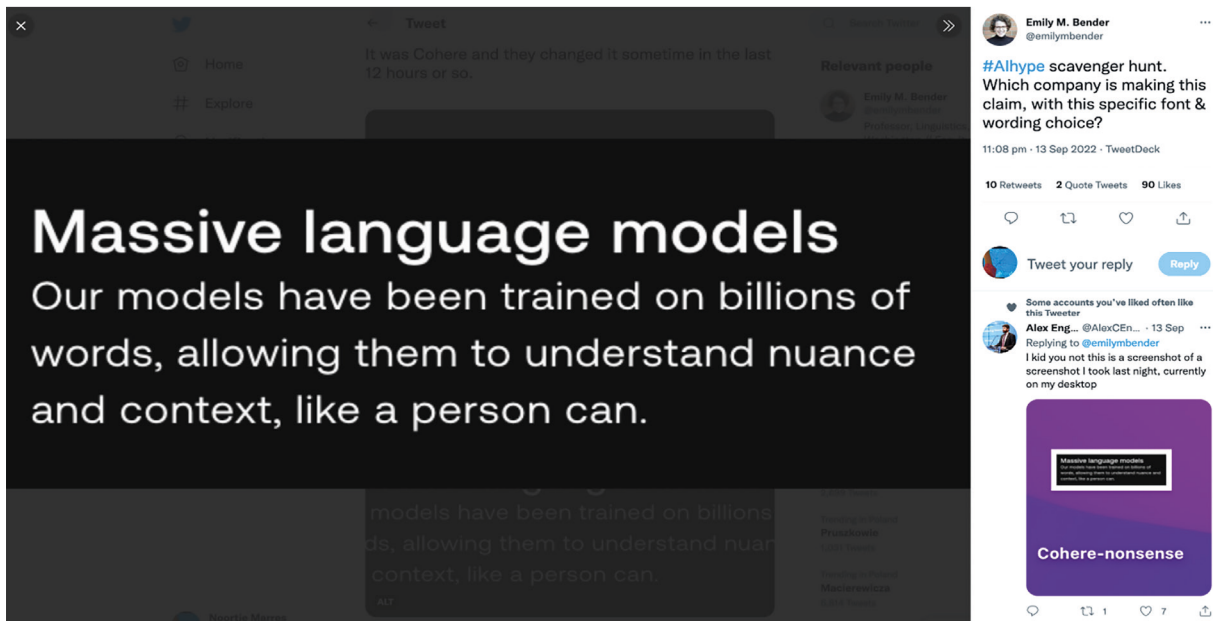


Figure 1: Emily Bender, Twitter quiz, 15 September 2022.

Our discussion is structured around the following three questions: First, we return to a classic critique that sociologists and anthropologists have levelled at AI, namely the claim that the ontology and epistemology underlying AI development is rationalist and individualist, and, as such, is marked by blind spots for the social, and in particular, situated or situational embedding of AI (Suchman, 1987, 2007; Star, 1989). We ask: Does the performance and evaluation of “machine intelligence” in contemporary instances of AI testing in so-called real-world environments continue to be marked by such a “social deficit”? Next, we delve into the issue of whether and how social studies of technology can account for AI testing in real-world settings in situational terms. Here we engage with the work of the French sociologist Louis Quéré, by addressing the question: What can we learn from today’s real-world testing of “AI” regarding the distribution of capacities between artefacts, environment and context in compute-intensive practices (Quéré, 1998)? And, thirdly, we ask: what does this tell us about possible tensions and alignments between different “definitions of the situation” assumed in social studies, engineering and computer science in relation to AI? Finally, we discuss the ramifications for our methodological commitment to “the situation” in the social study of AI: does it make sense for social studies of technology to continue to rely on the description of situations in undertaking the “re-specification” of machine intelligence?<sup>2</sup>

<sup>2</sup> Our initial conversation was interspersed with questions and comments from the audience, which have all been transcribed for present purposes, including a final reflection on turn-taking in (online) conversation.

### Question 1: Does the the new AI still rely on the bracketing of situations?

*Does the performance and evaluation of machine intelligence continue to demand the erasure of situations and the bracketing of social life?*

NM: To address this first question, I’d like to start with a particular challenge posed by the rise of learning-based, data-intensive AI. Especially challenging from the perspective of social studies of science and technology, I believe, are the sensational claims that have been made these last years regarding the capacities of these systems for “situational intelligence” and “contextual learning.” Here is a quote from computer scientist Percy Liang taken from his introduction to a Stanford University workshop on so-called large “foundation models”:

Foundation models [...] are based on a decades old idea, self-supervised learning, meaning that, based on lots of raw data, you make up predictive exercises [...] like weight training to develop the muscle for pattern recognition [...] Doing this at scale, results in the emergence of new capabilities, and one thing GPT can do is in context learning (generalization to new tasks) (Percy Liang, Centre for Research on Foundation Models, Stanford University, 23-24 August 2021).<sup>3</sup>

In the 1980s and 1990s science and technology studies (STS) and studies of AI in that field claimed that

<sup>3</sup> This introduction can be viewed here: <https://crfm.stanford.edu/workshop.html>

“context and situation” are precisely what automated systems are *not* able to take into account: this was the critique that Suchman (1987) and others leveled against older “expert systems,” and which Suchman (2008) claimed in relation to robotics, persist as “unreconstructed form of realism in roboticists’ constitution of the ‘situation’ [...]” even as “references to the situated nature of cognition and action have become ‘business as usual’ within AI research” (148–149). Today, it can seem that this argument does not quite obtain for AI, for a number of reasons. One is that the idea that the “New AI” is capable of “situational intelligence” or “context learning” is today being echoed by STS scholars. Take Harry Collins, who argues in his *Artificial Intelligence* (2018) that:

the problem for AI is how it can develop social abilities, because this would require the full embedding of AI in language speaking social communities in society. The problem of AI is the problem of engagement with social context. AI engagement with the Internet has resolved this to some extent (Collins, 2018: 162).

Collins, not unlike Percy Liang, appears to claim that training computational models on large volumes of Internet-derived data has solved AI’s problem with social context.

What I would like to emphasize, however, in relation to these kinds of claims, is how *incredibly selective* both a sociologist like Collins and a computer scientist like Liang are in their definitions of what counts as a relevant context or a relevant situation for AI to engage with. Context seems to be defined as the prior occurrence of a given utterance or interaction in textual or visual data, thereby excluding most of the features that sociologists regard as crucial attributes of situations (embodiment, materiality, co-presence). Furthermore, their notion of context seems to exclude the types of situations that the introduction of “AI” into society itself gives rise to. There are many examples of this, and many where we can see the process of “contextual learning” go *wrong*. Take the Uber delivery bot that rode into a crime scene (<https://incidentdatabase.ai/cite/350>) as well as other incidents in the AI incident data-base. Many of you will know the case of the racist online *chatbot Tay*, which became radicalized after it was trained by the 4 Chan community and gave us the spectacle of a racist online chat course (Sharma and Brooker, 2016), which subsequently got retrained and eventually dissolved.

These kinds of cases show us is that there is a lot of “context” which AI is not taking into account, and it suggests there is a lot of *perverse socialization*, and failed socialization, happening in “contextual” engagements of and with AI. In my view, the feminist critique of machine blindness to the world remains as relevant as it ever was. Indeed, sociologists, and social research more broadly, *have* drawn attention to these kinds of problematic interactions between artifact, en-

vironment and context, in cases like the perverse socialization of *Tay*. They’ve shown how the toxic online environments that appear to be the accepted social environment for large language model training (Bender et al., 2021) are co-producing monstrous forms of AI. They’ve also shown how the staging of a public situation involving a racist chatbot by Microsoft was conducive to the realization of that perversion (Neff and Nagy, 2016), a bad situation which resulted from the contextual blindness of AI developers.

However, sadly what has not yet followed from these sociological studies is widespread recognition of the challenge that these cases of perverse socialization pose to accepted definitions of what counts as “contextual learning” by AI in computer science. Instead of a critical grasp of the methodological and conceptual challenges that arise when computational systems operate in and as social life, what we often see instead is that these kinds of cases, like the racist chatbot, get framed as *problems of ethics*. This has the effect of placing the whole situational logic of how a bot becomes racist outside of the epistemological frame of AI development and research. And so “situational AI,” in the way that a sociologist would understand it, meaning AI systems that operate in social situations, still receives fairly little attention in the domain of AI development and research. And this is why I think it’s really important that social and cultural studies continue to insist that *situational enactments* of AI can demonstrate *methodological and conceptual problems* with AI.

I have one more quote to show you how easily the *erasure* of situational logics can happen in AI development and research. This is from an expert interview I recently conducted with a Connected Autonomous Vehicle-engineer, where I asked him about the complexity of the situations that automated vehicles encounter on the road. His response to this question was:

We believe quite strongly that the complexity in driving on the roads, is not in observing where the road ends and the pedestrian crossing starts and where the traffic lights are, these static tasks of identification have been solved for a long long time actually. The real challenge is modeling the behavior of other so called agents, because they’re not necessarily totally rational or perfect or identical.<sup>4</sup>

So it may appear that the kind of interactions that occur “in situ” are considered as part of the development of automated vehicle systems. But this engineer then went on to state that these situations can be *dealt with* by specifying the statistical properties of rule following. He went on:

<sup>4</sup> This quote is from an expert interview conducted on 28 May 2022, one of twelve interviews with UK-based Connected and Automated Vehicle (CAV) experts that I conducted during 2022–2023.



Figure 2: Alpha Go, as represented by Aja Huang (left), facing Lee Sedol (right). Source: press kit of Alpha Go. *The Movie*

... they still do follow rules of the road, but probably more fundamentally statistical properties, based on experience. And the way we learn these rules as children or as young adults is through observing the patterns of how vehicles move. And what we are implicitly learning is physics.<sup>5</sup>

In his view, it's really the *statistical models* in the *brains of individuals* that are the object of what an intelligent system has to simulate, if the aim is to navigate an on-the-road situation successfully. The fact that a situation is *interactionally* accomplished, between contingently positioned actors... this gets bracketed. So I think we do still need to be insisting on this understanding of situations as *accomplished* interactionally, and contextually, and this is what I believe contingencies are for – Philippe will have more to say on this.

PS: Yes, but to address the first question we gave ourselves for today, I want to start with two *conceptual remarks* and then introduce some *empirical examples* that I have been working on over last five years.

First off, I think it's still interesting to use the concept of "machine intelligence" because it reminds one of the *common ground* between what is often discussed as "AI," or "good old fashioned AI," top-down programming, rule-based, and so forth, *versus* current forms of "machine learning," and "deep learning" notably, where patterns are to be recognized in large datasets, and predictions and probabilities calculated on that basis, and so forth. But there's a commonality to this, and the term "machine intelligence" captures this quite nicely – a real *Iceberg* of assumptions, in fact,

so that one can reduce rule-based conduct to code, for example (which goes back to Turing at least, as does the notion of "machine intelligence"). D. Cardon et al. (2018) point out that current forms of "machine learning" are often pitched and promoted in terms of "AI," notwithstanding its programmatic purpose in the mid-1950s, namely to demote "machine learning" as viable research avenue in the field. The ensuing controversy, however, begs the question of its common ground, and how that common ground is artfully deployed – for example, in and as part of a technology demonstration (e.g., Sormani, 2022).

With respect to the question – "does the performance and evaluation of machine intelligence continue to demand the erasure of situations and the bracketing of social life?" – I would then *qualify* that question in the sense of pitching it not just as a yes/no-question, but by rephrasing it so as to include the demonstration and evaluation of "machine intelligence," and *how and why* that this technology demonstration and evaluation might bracket what the question identifies as "situations and social life"<sup>6</sup>.

Now, and with respect to *empirical examples*, my current focus is on "*edtech* in interaction," and how longstanding notions of "machine intelligence" are folded into participants' engagement with *edtech* in the classroom ("*edtech*" standing for *educational technology*, typically digital). That said, five years ago, I wasn't the only one to note the curious frenzy around

<sup>6</sup> Of course, our question could also be rephrased in similar terms: "how and why is the real-world set up today so that AI testing can happen within its remit, without creating a 'problematic' situation – an emergency, incident or accident – but instead maintaining a sense of normalcy, as an everyday scene?"

<sup>5</sup> See footnote 4

“New AI,” *inter alia* because it pitched “deep (machine) learning” in terms of “AI,” the term introduced in the mid-1950s precisely to demote “machine learning” as mentioned before (cf. Cardon et al., 2018). So the cases I started working on were public demonstrations of AI-labelled technologies, such as the *AlphaGo* exhibition match in 2016, before examining video footage from street trials (with “autonomous buses”) and now *edtech* in pedagogical experimentation (with various “educational robots”). Before in each case getting back to the *why* question, I’ll pause on the *how* question:

As a public demonstration of a sophisticated “AI” system, the *AlphaGo* exhibition match was broadcast in March 2016 from Seoul, South Korea, and pitched the system against Lee Sedol, a South Korean Go world champion at the time. Hence the question: *how* was the *AlphaGo* exhibition match staged in the first place, as a “symmetry spectacle” opposing *AlphaGo* and Lee Sedol (see Fig. 2).

Secondly, I also had become interested in a “smart shuttle” street trial, an interest that I have pursued together with Jakub Mlynář. On the website, the “smart shuttle” was presented as the “first intelligent bus in the world”<sup>7</sup>. Hence again the question: *how* were the streets – the bus, other vehicles, pedestrians, etc. – staged so that it could appear as such? And for the “Mars mission” as a classroom experiment, the case I have recently been concerned with: *how* is it staged so that “each robot can be controlled from Earth” by pupils in the classroom? I shall get back some of the technicalities later. In the meantime, let us address just how the mentioned demonstrations of technology bracket or even erase “situations of social life.”

One way of doing so is to take a closer look at how *contingencies* are managed and how they are managed for one version or other of “machine intelligence” to appear, similarly to specifying the “work of making an experiment work” (cf. Garfinkel, 2022). A starting point is provided by a distinction drawn by Garfinkel (2022), the provisional distinction between what he called “standing contingencies,” or we could call them “manifestly standing contingencies” tied to a particular practice – say, a Go exhibition match – *versus* “locally produced contingencies,” contingencies that present themselves in the course of the action (90–91, note 34), where there’s always this unanticipated or difficult-to-anticipate course that participants contribute to and are confronted with (aka “situations of social life”).

With respect to the *AlphaGo* exhibition match, there’s several features that can be described as “standing contingencies” tied to the exhibition match and its production as visibly a match of that kind. First, the game and commentator rooms are shown, not the control room, let alone the computational infrastructure that is necessary to have this exhibition match proceed.

A second example is, and this is from the documentary movie (Krieg & Kohs, 2017), the point at which the PR responsible says “we have to hide this somewhere,” – that is, the *AlphaGo* laptop computer – and she would put it underneath the desk that they – the professional player, Lee Sedol, and *AlphaGo*’s Aja Huang – will be using and playing on. And thirdly, and this is also from the documentary movie, Demis Hassabis, the CEO of the company behind the *AlphaGo* program, *DeepMind*, is shown to call Lee Sedol, the Go professional, to invite him to play the exhibition match against the program. And during the call, what appears behind Hassabis is the whiteboard, where everything is annotated for how the event should be staged – but that’s not something that is further elaborated upon in the movie either. Taken together, these aspects can be seen as an ensemble of standing contingencies with respect to this demonstration or exhibition match, and their local management has the “social situation” that the demonstration relies upon partly disappear from view.

But then in *reenacting* the thing, one encounters further contingencies – “locally produced contingencies” – that seem to have been generated and dealt with as part of the *scene shown directly*, such as when “Move 37” is played out, the professional commentator suggests “well, this is a move that nobody should be actually playing out at this moment.” But then in *replaying* the move, as an amateur player, I realized “oh, but this is actually a very powerful move,” *but in the most traditional terms*, a powerful move as described already in *Direction of Play*, an important Go book from 1979 (Kajiwara, 1979). So that’s how “power play” is achieved. Whilst “Move 37” was hesitantly qualified in Kajiwara’s terms also during the exhibition match, the initial implication on *DeepMind*’s website was less cautiously stated, the program, “*AlphaGo*,” being said to have “somehow taught the world completely new knowledge” (the commentary has been rephrased since). Again, the practical management of the contingencies involved, as already part of the scene, foregrounds the “machine intelligence” of the *AlphaGo* program, while backgrounding the local crafting of its scenic conditions.

I could list further contingencies for the other two examples, but I think in the interest of time I will not do that. But just stop briefly on the *why* question.

And that’s also a question that Phil Agre, a critical computer scientist, asked Garfinkel with respect to his interest in local contingency management and experimental scientific practice. “What are the contingencies for?” (see Garfinkel, 2022: 24; 39–55) What’s the point of listing them? And Garfinkel would include this question as a further contingency in the list, because the *why* question at times also becomes relevant for participants, sociologists included. And in this respect, I think there’s three matters that can be at least flagged. First, the issue of *accountability*: how do things appear? How are they shown? What are the consequences? Then, the question of how the *context* is “managed,”

<sup>7</sup> <https://www.postauto.ch/en/about-us-and-news/innovation/autonomous-driving>

given a situation's own momentum. And this is related to technology demonstrations and the point made by J. Lampel twenty years ago in a paper titled "Show-and-Tell: Product Demonstrations and Path Creation of Technological Change" (2001). The point is this: for a technology to be presented in a trustworthy way, you can't present it in too much detail. And obviously what is presented is carefully crafted. So that's a further aspect – that is, how critical inquiry is disabled or disfavored, while "commitment evaluation routines" (304) to the technology presented are foregrounded and favored. And then thirdly, and that's another classic issue, the risk of *reification*: if we omit local contingency management from sociological scrutiny, including dramaturgical uses of the "front and backstage" distinction, then we are at risk of reifying "machine intelligence" *ex nihilo*. Conversely, one gets a clearer sense of the "social situation" and its lively course, which are typically presupposed by and largely omitted from technology demonstrations. Back to Noortje.

**Question 2:  
What can be learned from the distribution of capacities between artefacts, environment, and context in AI testing situations?**

*What can we learn from today's real-world testing of "AI" regarding the distribution of capacities between artefacts, environment, and context in compute-intensive practices (Quéré, 1998)?*

NM: So let's take up this very general question of social studies, that of how researchers should distinguish between the role of "artefacts, environment, and context" in the study of computational practices, in which we can include the compute-intensive arrangements of "AI." To pose this question is to invoke the work of the French sociologist Louis Quéré who made the case for the importance of this distinction – between artefacts, environment, and context – and against conflating these terms in an article entitled "The still-neglected situation?" (Quéré, 1998).

I would like to highlight two reasons for why this question of "What belongs to the artefact? What belongs to the environment? And also what belongs to the context or situation?" is of special relevance to the study of contemporary "AI" and of AI testing.

First, it has often been pointed out that in demonstrations of AI, spectacular capacities – capacities suggestive of intelligence – that are attributed to the "machine itself," on closer scrutiny depend on active contributions from the machine's environment, including the humans that ensure its proper functioning, as Philippe just pointed out (but see already Latour, 1996). The study of demonstrations and tests of AI is marked by this analytic commitment: by studying "AI" testing in social environments, such as the testing of

self-driving cars in the street (see Fig. 3), we can investigate how the capacities of judgement and decision-making that are ascribed to "AI" are accomplished in situ. It is first of all within this context that the question posed by Quéré is relevant. When we study demos of AI, we can ask: what are the respective contributions of the artefact, the environment, and the context to the accomplishments of "AI" in each case?

Thus, in relation to a street trial of autonomous vehicles in Coventry city centre, as recorded by an accidental passer-by in November 2017 (Fig. 3), we can ask: what is the contribution of the safety cones that you see positioned next to the vehicle? What of the fencing? What of the labeling of the vehicle? And what of the safety guard you see standing in the background? In studying "AI" testing in the street – and there are many other cases, like facial recognition technologies that were trialed in train stations across the UK in recent years –, we are able to rely more on the social research method of field-based description, and less on the formal descriptions of "AI" technologies that are produced for public consumption by computer science, the tech sector and industry. Consider, for example, the public announcement of the Coventry street trial, which introduces this trial with the statement that it is "testing a number of features and most importantly seeking to investigate how self-driving vehicles interact with other road users."<sup>8</sup> You see here the other road users present in the form of pedestrians, with no fencing or cones or guards present in the image.

Perhaps unsurprisingly, when seen from the ground, an intelligent navigation system does not live up to the expectation that such promotional descriptions create, namely, that it is capable of coordinating its behaviour with other road users in situ. Instead, this "AI trial" is marked by confusion, and relies on all manner of props, with a fence plonked in the middle of the road and the passers-by not quite getting what is going on, and a guard managing traffic. *What does this rather confused situation tell us?* I think it tells us that when we consider AI trials, as they unfold as part of social life, we encounter a rather different type of situation, one that is clearly different from those stereotypical situations staged in "AI" trial demos, such as the demonstration of the Google call assistant called "Duplex" in

<sup>8</sup> The full quote reads: "The UK's largest trial to date of connected and autonomous vehicles technology on public roads explor[es] the benefits of having cars that can 'talk' to each other and their surroundings – with connected traffic lights, emergency vehicle warnings and emergency braking alerts. The vehicles rely on sensors to detect traffic, pedestrians and signals but have a human on board to react to emergencies. The trials are testing a number of features and most importantly seeking to investigate how self-driving vehicles interact with other road users." Tute, R. "Driverless vehicle testing on public roads hailed as landmark moment," 24 November 2017, see <http://www.infrastructure-intelligence.com/article/nov-2017/driverless-vehicle-testing-public-roads-hailed-landmark-moment>

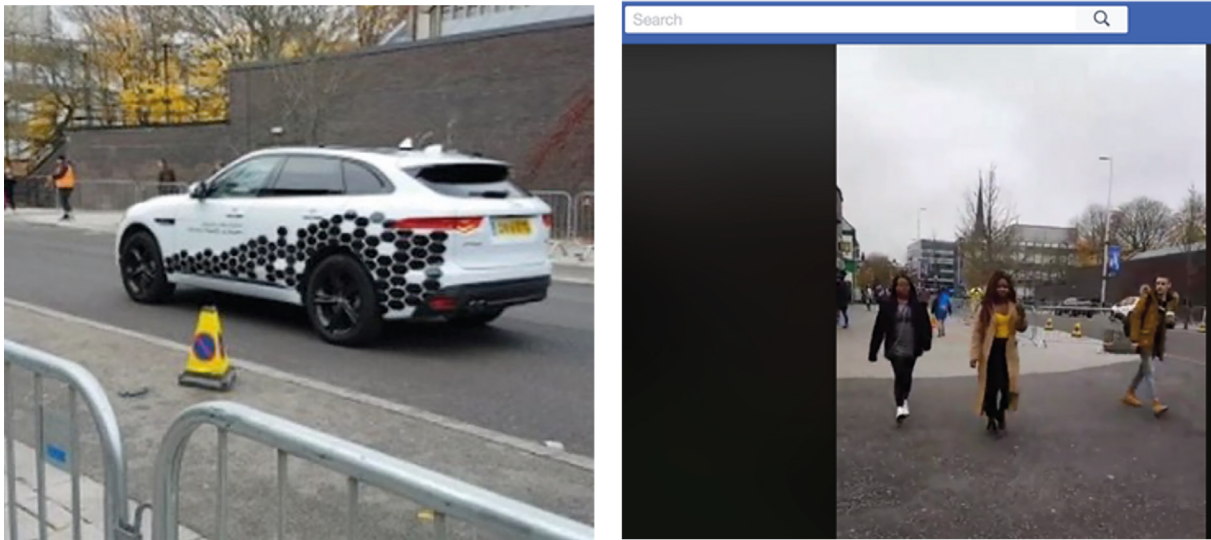


Figure 3: Coventry Autodrive autonomous vehicle trial, Coventry Telegraph, Facebook Live (November 2017)

May 2018.<sup>9</sup> This call assistant relies on a neural network and was supposedly capable of passing for a human, as was demonstrated through a supposedly live call to a hairdresser by Google's CEO on stage.

In the streets of Coventry, I did not find any systems trying to pass for humans or social actors. Instead I found a highly artificial situation, one in which it is not clear whether and how the technology is functioning, where actors seem rather disoriented, and there is a very heavy reliance on props, like fences. So I am making here the STS point discussed above, that the intelligence ascribed to the machine is in fact accomplished by a whole collective of actors present in the situation (fence, safety guard, cones, labels, etc.).

A testing situation like this also sheds light on what is perhaps a less obvious point, and this is that the introduction of "AI" into society entails modifications of environments in society, modifications which I believe trouble and to a degree undermine the very distinction between artefact and environment as Quéré makes it. The "performance" of this self-driving vehicle as an artefact is accomplished through *interventions and modifications of the setting*. In Figure 3 you can see fences that have been installed to facilitate the trial and there are also other, less visible modifications of the environment that took place as part of the Coventry trials: the installation of "road-side units," which included sensors and enabled communications between vehicles, as well as upgrades to road signage on the street surfaces so that these would be detectable by machine.

This is my *second* point: What the study of "AI" testing in situ highlights for me is that there may be "leakage" between these different, constitutive elements of social life: artefact, environment, and con-

text. Quéré (1998) frames these elements as distinct, arguing that an environment must be sharply distinguished from the context in which experience unfolds:

an environment in itself has neither axes nor directions since we are the ones who set them in different ways; these settings give rise to an 'enviroming experienced world'. [...] *It is the orientation of experience that gets one from the environment to the situation*, because situations come under the register of the organization of experience, which is not the case of environments. Someone who is disoriented is still in an environment. (288, emphasis added)

My proposition is that **this distinction between environment and context, and perhaps also artefact, is undergoing reconfiguration, socio-materially speaking as part of real-world testing of AI today**: as the roll-out of AI in society involves the insertion of compute- and data-intensive devices into the background of social life, and indeed the modification of infrastructural environments in society, artificial "intelligence" becomes literally the accomplishment of environmental modification. The artefact can't function without this modified environment, it is these computationally equipped environments that provide navigational guidance and enable vehicle-to-vehicle communication that play a crucial role in "providing orientation," to vehicles this time.

The distinction between artefact and environment, and context, I believe, is increasingly muddled in societies with AI, and that this is what the case of "AI" testing in society helps to make clear. The orientation of experience, which Quéré groups under the rubric of "situation," is very much what the curation of the environment is about: sensors in the setting provide guidance to vehicles, enabling them to navigate; fences guide the perception of the test. The situation, one

<sup>9</sup> See <https://www.youtube.com/watch?v=D5VN56jQMWM>

could say, is what *arises from* the embedding of compute-intensive systems into the socio-material environment. These systems transform the conditions for social routines in these settings. It changes the ways social life can unfold in them.

However, we should note that, as the environment is modified to provide orientation for machines, the social situation of the street is disrupted, and indeed, becomes dis-oriented. To refer back to the confusion noted above, and to put this somewhat dramatically: as long as the analytic focus rests on how the AI system manages to perform – how it orients itself and deals with ‘breakdown’ (aka ‘disruption’), in real-world tests like the one in Coventry –, we don’t really see how social life is very actively being broken down – dis-oriented – by the introduction of that very system. In this testing situation, the coordination of interactions between vehicles and pedestrians can no longer proceed “as normal”. The restrictive measures that ‘AI’ requires to function render routine interaction in the street impossible. And this schema, where the facilitation of machine orientation results in disorientation in the wider societal situation, is replicated at the larger scale. I am thinking of the profoundly disruptive and harmful effect[s] on the world that derive from continued investment in automobility.<sup>10</sup>

To sum up, “AI” testing invites re-examination of the formative distinctions in the social studies of technology, between artefact, environment and context. Indeed, in studying societies with AI, it becomes our task to investigate how capacities are being redistributed between artefact, environment, and context in the implementation of “AI” and the situational disorientation that ensues.

PS: Thank you, Noortje. From an ethnomethodological perspective, I have quite recently become interested in “AI” technologies, or AI-labeled technologies, including the use and development of them, as well as research on them – for example, but not exclusively, in the field of *edtech*. And, of course, I am not alone in this<sup>11</sup>.

<sup>10</sup> This is another reason why we should “de-naturalise” our account of AI, by focusing on situations of testing, or demo-ing. This empirical focus notably allows us to refuse the naturalistic fiction, and resultant deception that is introduced when the question is posed: how does “AI” deal with breakdown? (Collins, 2018). The methodical suggestion here is that AI must be deemed intelligent, if it does manage to deal with breakdown. But what really is the relation between AI and breakdown? A naturalistic approach has the effect of obscuring the disorientation, the confusion, the disruption and the breakdown that ensues as a consequence of the introduction of AI in social life. The social science that goes along with treating the artefact “as if” it were a social actor, is one that accepts as given the promotional descriptions above, instead of analysing the situation, of observing what actually happens when AI is introduced into social life, which a lot of the time is producing disorientation, and precipitating breakdown.

<sup>11</sup> Regarding *edtech* in interaction, I acknowledge instructive conversations with my Lausanne colleagues, including Marc Audétat, Julien Bugmann, Farinaz Fassa, Guillaume Guenat and

Regarding our point of discussion, question 2 above, I affiliate with the case for continuity made by you and David Stark. In the paper entitled “Put to the test” (Marres and Stark, 2020), you link “expert-led testing and social experimentation” (428, emphasis added). Separating them, in turn, “risks rendering invisible the testing situations that the sociology of testing should elucidate” (ibid.), but arguably any contemporary “ethnomethodology of experimentation” too. So that is a first point of convergence, a point of convergence which can be traced back to ethnomethodology’s longstanding interest in “heuristic handicaps.” Be they devised on purpose (e.g., via perceptual experiments), due to a physical disability or locally emerging, handicaps and problems of various kinds continue to offer an intriguing heuristic, a research path for explicating a social phenomenon in its own terms and techniques, *while* critiquing or at least avoiding “naturalistic fictions” (to use your term) from an impossibly detached position<sup>12</sup>.

However, and again, I also have two *conceptual caveats* with respect to the very formulation of this question, two reservations from an ethnomethodological point of view. And I might add a third one.

First, I don’t think that the “distribution of capacities,” and how they are attributed, is a good starting point for description, but rather one should start – as emphasized before – with *how contingencies are managed* in situ, and how they are *manifestly* managed, so that they can be described in the first place – if only to avoid “distributed essentialism” (Woolgar, 2004) in the service of actor identification. That is to say, an ethnomethodological description – a description that homes in on the everyday methods of practical activities, their particular intelligibility and its situated accomplishment – *ends* where a sociological model of agency explanation starts or might start, insofar as such a model *presupposes* what the description provides – a recognizable situation, a perceptible configuration, an unfolding interaction (in terms of which “actors” are identified, “capacities” distributed, “obstacles” spotted, and so forth).

The second reservation I have relates to the idea of “compute-intensive practices,” as constitutive of “computational artefacts” – be they programs, programs with sensors, or programs with sensors and actuators, as Johnson and Verdicchio (2017) distinguish them. Again, neither “compute-intensive practices” nor “computational artifacts” appear as a good starting point for description, insofar as they (as concepts) are too restricted. One risks missing the *cultural* artefacts that are involved, how they are produced *in situ*, and what *practices* constitute them, be they “computational” or other kinds of practices. In the slide that

Audrey Hostettler in particular, as we have become involved in “relocating machine intelligence.”

<sup>12</sup> On heuristic handicaps and their critical interest, see already Lynch et al. (1983).

Noortje showed just before (Fig. 3), we see first “a car,” or we see “pedestrians,” so we see them as a cultural artifact or embodied actors, and not as computational ones, to begin with. If they are to be considered as “sociotechnical assemblages” (Both, 2020) – “embodied” and “entangled,” “never pure,” “never alone,” – what kind of assemblages are they, how are they put to work, and how do their ‘workings’ play out *in situ*? And what then is the *situated praxeology* of a (if not the) “cultural life of machine learning” (Roberge and Castelle, 2021)?

A third reservation – or possible reservation – takes its cue from Quéré’s key distinction between “environment, context and situation” (1998: 243; emphasis added). I agree with you that AI systems, for them to operate as part of a street trial, require that the “environment is modified” and that this instrumental modification may prove problematic to, if not disruptive of, regular traffic. In that sense, the “social situation of the street is disrupted.” However, a disrupted situation remains a social situation. Quéré’s distinction, then, is useful to hint at the difference between a specific *context* or selective contextualization of a street environment, on the one hand, and how an everyday *situation* of traffic (whatever its participants, context or reframing) actually unfolds in its rich particulars, on the other. A “situation” may simply prove irreducible to any “environment” or “context” per se, as a “lost newcomer finds himself suddenly in the midst of a Mexico City traffic circle” (Sudnow, 1978: 30; emphasis added).

Now, let me get back to my three examples: *edtech* in “Mars mission” interaction, a “smart shuttle” making an impromptu stop, and *AlphaGo* on stage. If these are *empirical examples*, what are they examples of?

Before I mentioned the local management of practical contingencies as a phenomenon of ethnomethodological interest. How do participants set up, use, and interact with a particular device, system, or infrastructure so that its operation may be said to display “AI” capacities? How do they do so, recognizably so? And what are the contingencies – the “locally lived constraints” (Garfinkel, 2022: 23) – they encounter, contend with, and/or subvert in so doing? Each of the three mentioned cases offers an empirical answer to the raised questions, an answer which each time, through the contingencies encountered, casts into relief Quéré’s distinction between “context and situation.”

The *AlphaGo* exhibition match in 2016 led to this particular move in game two – move 37 –, which was commented upon as being a very special move by the “AI” system, having “somehow taught the world completely new knowledge.” This media announcement, as it was initially posted on the DeepMind website, encapsulates the “surprise moment” of the English-speaking game commentators as they first noticed *AlphaGo*’s “move 37” during the exhibition match. One of them, Michael Redmond, a top Go professional himself, was arguing against playing any move into the area (i.e., the area on the right-hand side of the Go board) into

which “move 37” had just been played by *AlphaGo*, the “AI” system whose moves were placed on the board by its lead programmer, Aja Huang, on stage. The discrepancy between the commentary and the move had the game commentators struggle to make sense of it. The discrepancy became their “locally lived constraint” for the ensuing commentary, not to mention Lee Sedol’s response move (as he confirmed after the game). In Quéré’s terms, the encountered contingency marks the difference between an expected, if projected, *context* and the unfolding *situation*, as actually dealt with.

A similar case could be observed during a street trial of the “smart shuttle,” where the starting point again was a locally encountered discrepancy, where something happened that shouldn’t or wasn’t expected by participants in the setting. In this case, we have a commentary by one of the passengers on the shuttle’s capacities, the passenger saying “it is very good how it does, going around things. I’m amazed that it goes through the narrow places so easily – yeah.” And at that moment, the bus stopped abruptly, and the operator commented on that stop as a recurring problem: “and here she [the van, *la navette* in French] does each time the same [thing] to us.” – that is, the van just stopped, without no apparent reason, at least for the operator. Again, the *context*, this time of a street trial, appears as different from the *situation*, as encountered and dealt with by participants.

And a third case, my current focus: Mars missions. In my home canton, these happen to be staged at school, where pupils are invited to program small mobile robots for them to move around and accomplish a mission on Mars, the Mars surface being staged at a technical university and shown to pupils via a YouTube link. Then, in the classroom, *some* of their comments would cut through the general enthusiasm: “but Sir, we’re not on Mars there!” (referring to the video stream), “we’ll be missing the competition” (referring to their sports day out), “I’d like to film” (referring to the camera equipment), and so forth. The list of contingencies encountered *in situ* could be extended. Again, it suggests how the particular *situation* proves irreducible to any given *context* or tentatively controlled environment (e.g., a “Mars mission” staged in the classroom)<sup>13</sup>.

So there are different kinds of contingencies that have to be managed, when different kinds of “machine intelligence” are staged, during different technology demonstrations, street trials, and/or pedagogical experiments. Of course, these demonstrations, trials, and experiments are also designed to have the contingencies *disappear*, as part of their practical management, management that makes them “work,” for machine intelligence to appear, be it as a rhetorical effect, a

<sup>13</sup> In turn, these locally encountered contingencies offer innumerable pedagogical opportunities, a point made by a colleague at the local teacher training university. For the programmatic argument, see Lynch (2022).

navigation requirement, or a pedagogical task. However, there are limits to this intended “disappearance” of mundane practicality. In one of my favourite quotes, Lucy Suchman puts the matter as follows: “lived practice inevitably exceeds the enframing moves of its own procedures of order production” (Suchman, 2007: 193). Okay, I think that’s all from me on this second question, and I believe it’s back to Noortje.

NM: That’s great, Philippe, I’ll make two quick points. First, the management of contingency, and also perhaps the undo-ability sometimes of contingent situations, I am very interested in this topic, in part because I am reading in social theory at the moment, including Agnes Heller’s *Can Modernity Survive?* (1990). This social theorist following Lukacs places great emphasis on the contingency of everyday life, as what is somehow distinctive of modernity: in modernity everyday life is experienced as *contingent*, it is a form of life in which routines and practices can be probed and challenged, indeed: tested, and modified because they are recognized to be contingent, and this is a key aspect of Heller’s understanding of modernity, and indeed also of her understanding of why we really need to work on ensuring it can survive. So, yes.

But I also think, and that’s my second point, that this question of contingency of the situation and the distribution of capacities within it are really closely connected. So one of the reasons I keep insisting on this question of how capacities are distributed or redistributed within a situation is that the way in which capacities get concentrated or consolidated within the artefact – alongside the elevation of the machine through the specification of its capacities as incredible, sophisticated, extraordinary... and all that – one of the effects of that is to make contingency disappear: it makes the machine appear as the necessary, inevitable executor of the task at hand – being extraordinary and thus irreplaceable, or this is one of the risks at the very least. While to note that capacities are distributed is to indicate that any distribution of capacities given in the situation is contingent, and may change. So I would say that to insist on contingency and insist on distributedness of capacity are perhaps not as much in opposition as you might have suggested.

PS: Thank you, Noortje. Let me briefly respond. I noticed that indeed you *start out* with reading Quéré (1998) *methodologically*, in the sense of getting or taking his question in terms of how researchers, and social scientists for short, *should* distinguish between artefacts, environment and context (if not between “context and situation”). And that was also his plea, his argument, since he thought they are too much lumped together, at least at his time of writing and the research situation that Quéré was commenting upon in the late 1990s, notably regarding workplace studies (in “human-computer interaction” and “computer-supported coopera-

tive work”) and object-interested sociology – “ANT.” On that count, I would agree also with the fact that, maybe instead of treating this as a methodological issue, we can look at it as a *phenomenon*: how was this done by researchers – sociologists, “AI” researchers, or participants indeed – in particular situations? To look at it as an empirical phenomenon, rather than quibble about methodology, not to mention “ontology.”

Now, this brings us back to the question of the “distribution of capacities,” and I wonder: is your argument about *multiple causalities*, in terms of how different capacities contribute to, say, an ongoing action? Perhaps this is too stark a formulation, but it allows me to mark a contrast with regard to what Quéré in my view was after. At least on his phenomenological side, his key interest doesn’t seem to have been how capacities can be distributed and attributed, but rather how a situation is intelligibly produced, as such and such – as a “whole,” as a Gestalt of a particular kind, which then only allows one to identify particular actors, in terms of a particular contextualization. So it is something that would come before the “distribution of capacities” and their attributions to different actors. But one doesn’t exclude – or does typically entail – the other. As you said, the political danger is that the way in which capacities are distributed has contingencies disappear, in tricky ways, although this danger might be the very purpose of successful engineering, at least in engineers’ typical terms.

CG: We have one question from the audience, by Erhard Schüttpelz (ES).

ES: Unfortunately, I have to leave to go teaching, so that it is why I thought I pose the first question now. Briefly put, from the outside, the differences between you are *minimal*. There is a perfect match, and there is a perfect match also in this question about how “AI” is staged, how it is propped up, and how it is filmed, and so on. So you can easily document your statements.

The terminology, of course, is the holy cow of each discipline. If a German starts a sentence with the word “Historically, ...”, the idea of the environment is the situation. [...] <sup>14</sup> On the other hand, the situation is also soaking up environmental motives in its course. So it is pretty arbitrary where you start and where you want to go. So a “situation” of course is also a word that I think has been misused a lot, too. So “we’ve got a situation here,” I feel, and the situation could also be messy, so [we move on]. [...] And also when Garfinkel’s “documentary method” chapter was defined by [...] “define the situation” through it, [he] wouldn’t get very far in [methodological] terms. So I stop here, because I need to leave.

Terminology is basically a discipline thing, and I just want to get to this aphorism – Can Modernity Survive?

<sup>14</sup> The somewhat sketchy transcription in this paragraph is due to the poor audio quality of the initial recording.

–, because I like that very much. But I think the answer is obvious, and was given by Latour years ago – in fact, thirty years ago: *Can we survive being modern?* If we want to survive, we cannot be modern (cf. Latour, 1993). Agnes Heller comes from a modernist school, so modernism must be saved. But I don't think we have to save anything from modernism. We have to survive by escaping modernism, after all. And don't think we have to keep anything modern, rather escape it so as to survive...

NM: That's fabulous, Erhard. (.) I have to laugh because, apparently, Philippe's and my attempt to be in disagreement hasn't been very successful.

PS: It's not over yet!

NM: If I can add one layer to that, we then maybe we also can connect it to the question of what modernity is. So one way one can set up the contrast – if one overdoes it, potentially – is to say, we have a political situation. We have a “political situation” in the sense that Andrew Barry (2001) uses that term, where there are fundamental disputes, and disagreements, and frictions in society which activate all kinds of structural problematics, and also issues to do with asymmetry and grievance. So, I am positing the situation of AI at a scale, and I want to posit it at a scale where we can appreciate it as a political situation.

I think that, in doing that, as social researchers, that it is very important that we mobilize the contextual and situational specificity of AI. So in a way, we are operating on two levels (and this something that I think we will also explore in the next two questions). In relation to modernity, what I think is so important about the work of Agnes Heller on contingency as constitutive of everyday life is her points – or at least she opens the possibility that for some among us – modernity *must* survive: if the contingency of everyday life becomes bracketed to the point that we cannot critically probe the routines and given forms of everyday life, then our positions – our positions and our orientation as working women, as actors who do not have a support machinery available to them in the way that self-driving vehicles have it – may simply not be tenable. For me the question of contingency and our need for it are connected to our positions as nontraditional subjects in the world. So, I think your question – *Can we survive to be modern?* – is really excellent and must be immediately posed after, and I'm going to ponder it, but this is why I can't give up the first one in taking up the second.

PS: Yeah. Maybe from my side on this question. Two things, and one thing I might have forgotten before, but obviously this emphasis on *contingencies* is related to the recent editions by Mike Lynch of Garfinkel's manuscripts from the 1980s on the sciences (cf. Garfinkel, 2022), and on how contingencies are dealt with in the laboratory, to “make experiments work” with all

the square quotes that one can put at around each of the words I just used. Now with respect to that Garfinkelian interest what I found interesting – in and as part of a reflexive ethnography (e.g., Sormani, 2019) – is to redo the phenomenon *and* lamentably fail in its redoing because, methodologically, it's a reminder of what the *critical* contingencies are that must be taken into account for, say, this or that device to appear “intelligent” on stage – it could also be an experiment, whatever. So that's one point also in a critique of interaction analysis, mostly starting from talk, because when technologies are used it is not only or primarily done through talk. Talk is part of another situation, if you will, with its locally relevant materialities and so forth.

The second point, also in listening to Noortje's comments on Agnes Heller's philosophical work on modernity, is in connection with efforts, notably by Melvin Pollner, in a posthumously published essay on how to connect Garfinkel with reflections on how modernity develops and unfolds (Pollner, 2012). The essay notably connects Max Weber's interest in bureaucracy and Garfinkel's interest in what he called “*formal analysis*,” and how that can be connected and also be put to use, in and as part of a *critical inquiry*, whose contours and content will vary, and to whose sharpening ethnomethodology may contribute.

### Discussion: Q & A Intermezzo

JS: It's just a question of clarification, Philippe, when you mentioned before that they were hiding sort of the computer on which the *AlphaGo* was to be presented. I was wondering *why*, maybe I just didn't get it, what was the reasoning behind it? Because I had the feeling that it was in order to keep or to present *AlphaGo* as a real player, so then it is a *real* actor, and as soon as we would see the hardware and the infrastructure, we wouldn't perceive it as a real player. So in terms of it's sort of like this uncanny sort of this uncanny valley, it's sort of a *reverse uncanny valley* sort of thing. Is that the direction you're going?

PS: Well, yes and no, I guess. Yes, in the sense that that was how the staging work was shown of this exhibition match – well, ironically or incidentally, by this documentary movie, which was produced by DeepMind itself. So it's also part obviously of the PR campaign around the *AlphaGo* program. So it was something I noticed in watching this movie in terms of how the staff seemed to be building the stage for the exhibition match to take place, and indeed to have it take place in a format that resembles a regular professional Go match, where you have a game room with two players and a commentator room. Similarly, in the exhibition match, there were also two players, while one of them had a screen to look at what move the program was suggesting, and a commentator room as well. But the technological infrastructure, the processes behind, etc., this was not

shown, neither the control room – I mean, at least not in the livestreams. So it was an empirical observation. Now how does that relate to the discussion about the uncanny valley, which is more connected to robotics, I wouldn't have an immediate response to that – hence my no, if you will. What comes more directly to mind, if paradoxically, is the intricate genealogy from stage magic to technology demonstrations (for a fascinating paper, see Smith, 2015).

CG: I would like to ask Noortje who was criticizing the capacities for “contextual learning,” whether this is basically a plea for developers to include this kind of accomplishment in contextual understanding and, if so, how do we then draw, how can we then draw the boundaries of what counts as a *relevant* context? and what counts as kind of the *limit* of this kind of distributed accomplishment? Because, you know, we have been studying that here in the context of the CRC for the last almost eight years, and it's so difficult to really draw the boundaries of what we consider medium-specific practices as being accomplished in a distributed way. What do we take into account in this distributedness? And so what do you take into account in the context? If we don't look at it from a pure developers' perspective, but also from taking a wider angle, and taking a more- taking *social concerns* also into account, and what would be the desired context [which] should be taken into account?

NM: Great question, Carolin. First of all, what I'm criticizing is the way in which, when AI situations themselves turn into disruptions, and when the harmful impacts of AI are situationally demonstrated, that this then gets framed as an ethical problem with the *application* of AI, as not having to do with the methodology, the epistemology of machine learning itself. So I'm first of all making a plea that we that we don't go along automatically with the strict internalism and externalism imposed by computer science, as to what is part of the science of AI proper, and what belongs merely to do the domain of application. I am criticizing the kind of highly asymmetrical *framings*, where, when we have an AI interacting politely with a hairdresser, we call this a demonstration of the capabilities of AI, and that, when we have a bot being radicalized by 4-Chan communities, that we call that a demonstration of the ethical dangers of AI. So that the stereotypical situation is one where the system displays its capacities, and the disruptive, harmful situation amounts to an ethical problem with application.

Instead, in *both* cases, we have an AI *situation* where AI is being *interactionally* accomplished, in the first case through a scripted demo, and in the second through less scripted though not unexpected user interventions, and both of these situations are part of AI. Of course, I'm not saying that all computer science should be working on interactions and situations and the situational ac-

complishment of machine intelligence, but that at least they get recognized as part of AI. To accept that the social aspect of something technical is part of the socio-technical phenomenon, that you can't just cherry-pick the contexts that are directly relevant from a computational development point of view.

But, of course, the *issue of selectivity* of which situations we engage with is, of course, you know, I can't pretend that it isn't also very much a problem for *us* in conducting social and cultural studies.

I don't want to banalize this problem of selectivity, but what I think should be our goal is to bring to the fore all kinds of situations, including ones where it may not be instantly clear *what the relation of relevance* is, because, when you look at mechanisms like “algorithmic impact assessments,” where the case is being made today for “AI” impact assessments, you often get a very sort of *instrumental definition*, almost mechanical, meaning, a-situational definitions: Who is the “community” affected? Who are the “effective stakeholders” who need to be involved in the design and the consultation on the design, *as if* one can always determine what the relevant community is, from the standpoint of the process of technology development. AI impact assessment involves recognition that AI can give rise to societal situations, and this is clearly an important step in the right direction. But methodologically speaking, the approach risks replicating an a-situational methodology of “stakeholdering,” and I think this is yet another important reason to champion and develop situational perspectives on AI. From a situational perspective, actors are likely to turn out to be relevant in ways that no one in the design process would have been able to recognize, let alone anticipate.

For instance, in the Coventry situation discussed above, citizen journalists played a key role in publicizing this trial and showcasing the disorientation it gave rise to. But it's not the case that Jaguar Land Rover, the local automotive company, recognizes citizen journalism in Coventry as a relevant actor, or user group. Yet they really played a crucial role. And so this kind of relevance relations as they emerge from contexts of tests, testing – those are the ones that I really want to emphasise.

Of course, we know these systems, learning systems are very dynamic, their objectives are always being optimized, they're always changing. So the very idea that you can just offer a static definition of who belongs to the community being affected, and once we have included them we're good, it just doesn't work methodologically speaking, even if ethically, it is certainly better than nothing. I am touching here on Louise Amoore's (2020) work, her critique of *source accountability* in Cloud Ethics. This book precisely goes in this direction of critiquing mechanical definitions of how relevance relations get established. If AI is beyond mechanical reason, surely we should move beyond mechanical conceptions of relevance.

PS: With respect to *situation*, rather than context, one aspect that Quéré (1998) also emphasizes is the *temporal dimension* there is to this – well, maybe when coming from a social science perspective, or at least in qualitative social sciences, one may have a tendency to *reduce* situations to interactions – but I think that’s wrong, and I’m not the only one, obviously, but I think Quéré’s paper is a nice reminder that situations could be anything from a “minimally complete conversation” like “hello, hello” (see Sacks, 1992) to, say, a revolutionary development, that there’s – in addition to the spatial dimension – a *temporal* dimension to situations so that one doesn’t know in what direction a situation will lead, also in terms of its eventual outcome and the narrative structures to which it will lend itself. Conversely, Quéré points out that “contexts” lack this temporal dimension of situations.

NM: Yeah, that’s important. I do think though that the *interactional* is quite key, in that it brings to the fore *differently* positioned agents, among whom coordination is somehow required, to even *do* whatever is being done. Shall we get to our next question, and then we can just open it up again?

## Conversation (ctd)

### Question 3:

**What are the possible tensions and alignments between different definitions of the situations assumed in social studies, engineering and computer science?**

CG: So the third question is: “*What does this – your conversation so far – tell us about possible tensions and alignments between different definitions of the situations assumed in social studies, engineering and computer science?*” And I think the last questions that we discussed were pointing towards that as well. But who would like to start?

NM: Let me say a bit about how I understand different definitions of the situation in computer science, engineering, in social studies and in other fields.

First, “situation” has been defined in many ways in sociology and social and cultural studies, so I just want to flag a few aspects of it. I always really liked the definition that Erving Goffman gives of “situation,” where he asks “what is going on here?”, question mark (Goffman, 1974: 8). Those moments when it isn’t clear what the answer is to the question “what’s going on here?”, those are the ones to watch out for if you are looking for situations. More precisely, those moments where answers are multiplying to the question, “what’s going on here?” – and it isn’t immediately apparent which ones are correct and which wrong, that’s when we have a situation, says Goffman.

Luc Boltanski and Laurent Thevenot (1999) offer another sociological definition of the situation: we are dealing with a situation when it is no longer possible to go on in the usual way. Here, we get a definition of the situation that highlights its *testing* aspect.

Finally, I already alluded to Andrew Barry’s (2001) definition of the situation, where he points to those moments in which there is friction, disagreement, dispute, and which is accompanied by the drawing together of underlying forces, societal asymmetries, and grievances. So there’s a dispute, the dispute is perhaps the surface of the situation, but one must also include in the situation all the sorts of underlying structural tensions, and possible inequalities, that come to expression, however partially, in friction and dispute.

What these different definitions, of what a situation is, share is their affirmation of *constitutive ambiguity*, *ambiguity as constitutive of social life*. Situations are constitutively ambiguous – “What’s going on here?”, “In what way can we go on, if we can’t go on in the usual way?” I think that this contrasts strongly with how – through recent conversations with automated vehicle engineers – I have understood that they define situations. In systems development, the term that I keep encountering time and again in how engineers refer to the social world, is that of “*scenarios*.” Automated vehicles need to be able to deal with different scenarios: a kid on a bike or a van that has a birthday cake dropping out of the back, etc. Now, this notion of the scenario, and the idea that to effectively evaluate an AI-based system, *you need to have access to a wide range of scenarios, this way of phrasing the challenge of AI in society really does not acknowledge – I feel, sufficiently – this problem of constitutive ambiguity, because scenarios within engineering discourse seem to be appreciated precisely because of their ability to offer formalizable, determinate descriptions of the real-world.*

The whole advantage of the scenario, from a technical point of view, is that it is formalizable in a way that enables *disambiguation* of every aspect, so the real-world can be fully specified in its relevant aspects. Indeed, I often get the sense that the suggestion that we should somehow allow for ambiguity if the aim is to grasp real-world situations, from an engineering point of view, tends to be taken as an indication of a lack of rigour, that “one does not get the requirements of systems development.”

So I do believe there is an almost *essential tension* between sociological and engineering framings of real-world situations involving AI. As to how to deal with that, of course, to an extent we can rely on a *division of labor between engineering and social science*. For me “constitutive ambiguity” is a critical aspect of social reality, and to an extent this tension can be addressed by saying: Let me work on constitutive ambiguities in situations involving AI. Let the specialists in verification and validation of autonomous systems focus on the development and selection of determinate “scenarios.” But the problem in accepting such a division of labour,

is that you *disarticulate* the methodological and normative issues at stake in this disagreement between a real-world that is fully formalizable and one that is defined by constitutive ambiguity. There's a risk that you just contribute to *disarticulating AI as a situation*.

This risk is to an extent clarified in the examples that I showed of so-called intelligent vehicle testing in Coventry. One of the things that I think is profoundly *problematic, and recognizable in this case*, is that the social environment is being transformed as part of the implementation of real-world tests of intelligent technology. All kinds of monitoring equipment are being installed, roads are getting fenced off. On a different, larger scale, we know the "digital highway," as an infrastructure is being rolled out, which is leading to tighter surveillance, and the creation of different forms of management and control of life in the street. This has all kinds of consequences for which situations are possible, which situations *can* arise in a society that includes AI.

The real-world testing of AI, in many ways, part of a *profound* transformation of the infrastructural environment in society. To engage with that political situation – which includes the question "do we even want to invite these machines into our societies?" – as well as the question "what kind of political processes are required to arbitrate on this question in a way that has public legitimacy?" All of this, I believe, really requires that we do *not* just look for an easy, pragmatic settlement of the conflict between the approaches to socio-technical innovation implied in the notions of scenario and situation.

I believe that we have to *problematize* that so much analytic attention in the social and cultural studies of AI continues to be focused on the question "how these systems work?" how do machine learning systems work, rather than: "how do machine learning-based systems *not* work?" Instead of only asking how these machines deal with disruption, our job as social scientists should surely be, first all, to investigate whether and how these systems engender breakdowns and disruptions of how *social life* is organized? How is the introduction of intelligent systems *actively* producing or enabling disruption or even breakdown of sociality in everyday environments? Take the "digital road," which is a profoundly restrictive space where no, you won't find a bike there, you won't find variations of speed there – it's an atomistically managed setting. To call into question those machine-centric definitions of the situation that suggest that it's somehow all about what these machines are capable of – to challenge that approach, that should be one of our contributions.

PS: Thank you, Noortje. Maybe Erhard Schüttpelz was indeed right, we have many things in common that we might not have thought of, before this conversation this afternoon, which is a nice outcome as well.

With regard to *tensions* between different situations, regarding how engineers might be interested in them

in terms of scenario *versus* social scientists' interest in how they unfold in an ordinary or as part of an ordinary encounter of everyday life, in thinking about the question I was asking myself whether it might not be interesting, or helpful again, to just distinguish – maybe in Latourian terms – between "readymade situations," or *assuming* a readymade situation, *versus* studying "situations in the making" as they unfold, etc., and obviously with the different actors involved and so forth.

A "readymade situation" could then be described as a stable *environment* or something that is already stabilized in terms of a *scenario*, predefining a particular context for a testing situation. On the other hand, the "situation in the making" points or hints more at the process of stabilization and contextualization, and how things get turned into other things. Well, if we take those three things that Quéré (1998) was interested in – "environment," "context," and "situation" –, we would have the *situation* as the "wild animal," through a particular management of the *context*, which would be the "cage," put into the "zoo," a managed *environment* or a controlled environment. What happens to the "artifact," and AI as an artifact in particular, in this analogy? From an STS perspective, a key advantage is to avoid the presupposition of an "autonomous agent," as an independent variable of sorts, and instead describe its dependency on unfolding situations, projected contexts, transformed environments. From an ethnomethodological perspective, the ethnographic focus would be on the encountered situation, as both the projected context and transformed environment will be determined *in situ*, at least with respect to how they become part of a social interaction or practical activity – through reconfiguration, if not subversion. The same holds for "AI" as a scenic feature of the examined situation.

*Empirically* speaking, where could one take it from there? One pending question, at least from the present perspective, is: how are different kinds of accountability – practical, technical, normative – articulated, and how is the articulation dealt with – as a connection and expression? Because from an ethnomethodological point of view, any situation is *accountable*, or any action is already *produced* as an accountable one – that is, as a distinctively "observable and documentable" one. But that doesn't determine yet: well, do we take this in a technical direction? Do we take it in a political one? And how is *that* actually done, by participants themselves to begin with? What kind of "politics of technology" is involved in either case? In other words: if an "AI" system involves a "performance," which kind of performance – or kinds of performances or aspects of performance – are we talking about? How does it, how do they, happen to be articulated? And who is "we" in this or that case? "Machine-centric definitions of the situation," to use Noortje's expression, address as much as they beg these sorts of questions.

Thirdly, and to get back to my three examples in the shortest way possible and also linking up with at least

how I started out this afternoon, there is something that struck me in these examples, namely: the *incongruity* that appears in the staging.

For example, the “smart shuttle” is first commented upon as “going smoothly around the town” and, all of a sudden, it *stops* – triggering the opposite commentary, “oh no, it’s always the same problem here.” In the *AlphaGo* episode, the program – assisted by its lead programmer – plays into the area “where it shouldn’t play into” at the very moment that the professional game commentator says this, or in the case of the “Mars mission” at school, one further incongruity was the following: “well, we stage this, but we don’t have any time for it” – at least not in terms of the existing curriculum. The listed incongruities exhibit the staging as such, the technology demonstration as a multifaceted performance. How then is the resulting situation dealt with? Who benefits from this, under what terms (e.g., a “pilot study”) and in what respect? How is actually a *division of labor* folded into the development and use of these infrastructures and these devices – something we didn’t talk that much about today, actually (although Noortje pointed out the political incongruity of a strict engineer/sociologist divide). And, conversely, what is the technology doing in and to the division of labor between different professional constituencies involved – engineers and teachers, for example, or also sociologists, obviously. *That* is not exactly what I wanted to emphasize now, but I guess that’s part of the situation.

CG: Noortje, any immediate responses?

NM: Yeah, quickly. I am especially interested in those *dramatizations of accountability* that are often *not* being accomplished. While we often find in AI testing situations rich enactments of accountability, these rarely translate, I feel, into effective articulations of *questions of political accountability*.

To give an example, Jack Stilgoe, who also works on autonomous vehicles in the UK, mentioned the case of an unmanned vehicle in San Francisco. A policeman walked up to a vehicle that was parked in the wrong place, but there was no driver in it. So you see this video of the policeman walking around this vehicle, but there’s no one to fine. There’s also no phone number.

On the one hand, you can say, this is a good situational dramatization of a larger problem of public accountability, which also highlights the problem of the immense privilege offered to these companies to run whatever trial they want to run in the street, in public space. The absence of any effective *subjectification of responsibility on the part of technological innovators*, it’s all there in that situation. But, at the same time, when I then look at, for instance, the debates about, let’s say, the AI Regulation Act in the EU, I do not find there a public staging of situational accountability either. I see AI being staged there as a *regulatory problem*.

The *efficacy* of staging AI as a *political situation*, that is something that in Europe, I think, requires more attention in the sense that there’s much invested in the *regulatory staging* of AI, *but* how does that *really* connect with how AI presents as a problematic situation – in social and public life? I think that’s often not considered in rigorous ways, but mostly by relying on stereotypes. Conversely, a reflexive interest in “performance(s)” may challenge such stereotypes, while recovering the “politics of infrastructure” that their routine use enacts and relies upon.

### Discussion (ctd): Final Q & A

CG: Thanks, so let’s open up then for questions from here in the audience and the audience online. I would like to *start* by coming back to one sentence I noted down: “in a situation where people cannot go on in the normal way,” and I was wondering: okay, do we then have to draw boundaries between *situations* and *moments of crisis*? This also is a key point of debate as to how various crises are clustered to different dimensions of social and technical life at the moment. *Is the definition of the situation that you have, is that blending into something we understand as a crisis?* [...] Or is that not relevant, because a crisis is such a specific situation?

NM: when you repeat the line “we can’t go on in a normal way,” I’m reminded of all those impossible situations which are cataloged in abundance online, such as the *bot* on the phone that doesn’t respond or there is some kind of *looping* in automated processes, where you just cannot get out of the loop, and it all becomes absurd.

Part of the issue – and this is broader problem, not just in relation to AI, but for all sorts of automated processes – is that there are so *many* critical moments. The abundance and the excess of moments where “we can’t go on in the normal way” is so great, and also so often occurring, and that has *shocking implications*.

You may know the case of how in the UK an automated system for processing the English language test results for asylum seekers resulted in false flag identifications of some of the candidates as *frauds*. This led to them actually being, their asylum claims being denied. There are *so many* situations where people can’t go on in the usual way in societies undergoing automation. This touches on the question that Philippe asked: how do you connect these practical situations to a definition of *crisis* that can gain some efficacy as a public definition of crisis, a sense that it is “not possible to go on in this way.” It’s a huge political challenge, but I think it’s also intellectually, and analytically, a question of how we make those scalar inferences between undoable situations and crisis moments? Though this may be a particularly British perspective.

PS: With respect to this point, and it relates to Carolin's question, I think, because I had the same concern or a similar one when I listened to your definition of situation as covering those where there's a *problem* involved and which have a "constitutive ambiguity," I think that was the term you used, because it's not quite clear how it might unfold or how it should unfold, indeed. But I guess even from a sociological or an everyday perspective, this is *one* kind of situation, a *problematic* one. I mean, there's all sorts of situation. In turn, the key move is to home in on at least *one* situation to figure out how it plays out, and how contextual and environmental aspects get determined in its course (which was also Goffman's concern in his 1964 essay with respect to structural models of language and social order).

So maybe analogously to Carolin's question – where she asked about the connection between, well, what's the difference between the *situation* you're speaking of and *crisis* – maybe one could also think of the situation as part of what kind of *inquiry* in Dewey's (1938) sense, given that that the testing of AI is part of – well, I hope, in some sense – a research and development process. In Dewey's understanding of inquiry, one aspect is that it leads from an *unclear* situation to clarification of "what is going on." So maybe, with our focus this afternoon on situation, we are at risk of forgetting what *inquiry* a situation is part of, or even what *society* different kinds of inquiries contribute to and situations become part of – in short, the situation as part of something else. In this respect, the contrast between prolonged inquiry and polemic shortcut in media discourse is of particular interest, too (cf. Quéré and Terzi, 2015).

With respect to *scenarios*, and this is just an empirical side note: we find them in AI testing of autonomous vehicles, we find them in *teaching* with educational technology. "Oh, we need a scenario!" – that is, for how can we use this tablet, robot, or other device for engaging with pupils, because obviously we don't want to start off – in primary school, that is – with some algorithm or the principles of computing, because that's not something that they are supposed to be able, directly willing or interested to latch on to. So we need a scenario.

CG: To what extent do we have to connect the question of the situation to the question of *sensing*? because we have been talking about AI, about the decisions made based on the situation, but for that we need also to, first of all, *sense the situation*. And that reminds me of the research situation we had last year.

We were looking into Waymo open data sets – these are open data sets for autonomous driving – and they were also labelled as consisting of "interesting situations," situations in which decision making for autonomous driving had to be performed. Then, there was a challenge for various researchers to resolve certain tasks, and we started to look into that data and what kind of object, or what kind of sense making of the situation, was provided in the data. And in the Waymo data, various ac-

tors that constitute a situation – such as pedestrians, cyclists, cars, etc. – were detected. Because we couldn't look into the data directly, we had to do a work-around, and in that work-around we could also look at the data through the Google vision APIs, the machine learning technology that allows to classify a large amount of images.

That was just fine for something very *different* for people [when] facing each other, with space or without space, people standing in a certain direction. So suddenly you would be exposed to very different kinds of categorizations of *what* constitutes this situation on the *image*, which are also down to questions of *sensing* capabilities of recognizing something, but also of categorizing. So: *where do you see the role of sensing here?*

NM: I think it's really hard. Probably, I find it hard because of the volume of scenarios involved. In the situations I've studied, it's about databases with thousands and thousands and thousands of scenarios. And, of course, one question can be, what is the ontology that is being presupposed? You could try to map out what is the world from the perspective of these entities that feature, and that could be interesting. But where I would want to make a start is to conduct interviews with engineers, and also more reflection on *how sensing is making a difference*. In London there's a lot of *irritation* about the degree of digital surveillance and data capture in test environments. Many of the most used test streets, at least in Greenwich, they feel rather like ghost zones: people don't go there much. So: to not take the sensors, the "sensing purity" as a perceptual device, but also look at it performatively. In any case, it would be great to learn more about the work you've been doing with scenarios, because that sounds really super interesting.

CG: Marcus [Burkhardt]?

MB: My question is as follows: why is it that all these issues, which are considered as problems and problematic with respect to "artificial intelligence," typically boil down to *anecdotal accidents*? Why is it the case that you raise just anecdotal evidence for these problematic situations? And might there be something *immanent* in this kind of technology, which generates problems, more so than scenarios, with stabilizing accounts? What might be drawn from this? [...] And how can we engage with and entangle these technologies in social situations and society as a whole, to define situations that we can contest and challenge?

NM: That's a really excellent question. I have two things to say.

The first is that within public policy debates, there is a relative lack of interest in *how accountability problems arise situationally*. The article "The still-neglected situation" (Quéré, 1998) touches upon what we're discussing here.

But what I also hear you point out is that, if we want to approach the situation in social, cultural studies – there’s an interesting book that just came out, entitled *Democratic Situations*, edited by Andreas Birnbak and Irina Papazu (2022) –, if we want to talk about democratic situations, then we need to work actively on this problem of the *anecdotal* as the seemingly only viable format for attending to situations. To posit AI testing as a situation *at scale*, and to ask, what are our methodologies for doing so, this is another way of connecting with Carolin’s question of how tools of data analytics can be deployed as part of this kind of studies.

CG: Philippe?

PS: With respect to Marcus Burkhardt’s question or observation, there’s some reflection in Langdon Winner’s 1977 *Autonomous Technology* book, which I think was his thesis at the time. In it, he observed that the focus on the anecdotal, the event or the media event of what happens, is also a sign of the *small opportunities for citizen participation* that are presented to, say, members of society with respect to modern technology.

But on the other hand, also he mentions in a phrase that, well, and here I quote, “the foundations of technological society are less reliable than some had hoped” (296). So, there is this interesting tension between on the one hand being entrapped in the “*société du spectacle*” versus this entrapment actually expresses something about the society we’re part of in terms of AI technology, precisely because a regularly vested interest might be to *reduce* possibilities for participation on a technological level more “upstream,” and not just to have the readymade scenario presented and the occasions on which it runs afoul.

CG: Thank you. Would you like to make a final statement?

PS: Thank you. Yes, I wish to make one [laughing]. The fourth question asks: “*Does it make sense for social studies of science and technology to continue to rely on the description of situations in undertaking the ‘respecification’ of machine intelligence?*” I would rather ask: “Does it make sense for social studies of science and technology *not* to rely on the description of situations?”, and I would say *no*. I mean, it makes sense to look at situations, but I think we agree on that?

NM: Here’s to more testing situations! PS: Cheers!

CG [et al.]: Thanks a lot to everyone. Bye, bye bye. Thank you. Bye bye.

[Zoom]: Recording stopped.

## Conclusion: pending issues and research prospects

*So, do we have situation?*

No, insofar as the “situational deficit” in how AI is framed, implemented and debated is still with us, contrary to the claims that machine-learning based systems are capable of contextual learning.

Yes, in that the introduction of AI into social life is giving rise to critical moments, and public and political situations, that remain under-specified in public discourse.

*So, what’s next?*

One move is to recover, and eventually redefine, the “testing” situations of AI, while both embedding and enlarging the notions of experiment and experimentation folded into the process. The preceding conversation alluded to first stabs in that direction, others have been underway for some time (e.g., Bogusz, 2022; Born and Barry, 2013; Marres et al., 2018) or remain still to be articulated, especially with respect to “AI” and its contemporary variants of machine learning.

*And what about turn-taking in (online) conversation?*

This working paper is based on an “augmented transcript” of the conversation recorded between the authors on the 25<sup>th</sup> of May 2022, as part of the Lecture Series *Testing Infrastructures* at the University of Siegen, Germany. On the one hand, the transcript gives readers a sense of the question-answer pairs through which turns at talk were allocated, both in response to the agenda-setting questions and in the ensuing Q&A exchanges. On the other hand, and for the purposes of this working paper, the transcript was elaborated, or “augmented,” by the authors as stated in the introduction, thus allowing them to complement, nuance or elaborate on the initial conversation.

In that sense, the working paper is an invitation to further elaboration and critical exchange, both on- and offline.

## References

- Amoore, L. 2020. *Cloud ethics: Algorithms and the attributes of ourselves and others*. Durham, UK: Duke University Press.
- Barry, A. 2001. *Political Machines: Governing a Technological Society*. New York, NY: Athlone.
- Bender, E. M., T. Gebru, A. McMillan-Major, and S. Shmitchell. March 2021. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?". *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*: 610–623.
- Birkbak, A., and I. Papazu (Eds.). 2022. *Democratic Situations*. Manchester, UK: Mattering Press.
- Bogusz, T. 2022. *Experimentalism and Sociology. From Crisis to Experience*. Cham: Springer.
- Boltanski, L., and L. Thévenot. 1999. "The sociology of critical capacity". *European journal of social theory* 2(3): 359–377.
- Bommasani, R., D. A. Hudson, E. Adeli, R. Altman, S. Arora, S. von Arx, and M. S. Bernstein et al. 2021. "On the opportunities and risks of foundation models". *arXiv preprint*. arXiv:2108.07258.
- Born, G., and A. Barry. 2013. "Art-Science: From Public Understanding to Public Experiment". In *Interdisciplinarity: Reconfigurations of the Social and Natural Sciences*, edited by A. Barry, and G. Born, 247–272. London; New York, NY: Routledge.
- Both, G. 2020. *Keeping Autonomous Driving Alive: An Ethnography of Visions, Masculinity and Fragility*. Opladen: Budrich Academic Press.
- Cardon, D., J.-P. Cointet, and A. Mazière. 2018. "La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle". [Translated as "Neurons Spike Back"] *Réseaux* 5(211): 173–220.
- Collins, H. 2018. *Artificial Intelligence: Against Humanity's Surrender to Computers*. Cambridge, MA: Polity Press.
- Dewey, J. 1938. *Logic: The Theory of Inquiry*. New York, NY: Henry Holt & Co.
- Garfinkel, H. 2022. *Harold Garfinkel: Studies of Work in the Sciences*, edited with an introduction by M. Lynch. London and New York, NY: Routledge.
- Goffman, E. 1964. "The Neglected Situation". *American Anthropologist* 66(6): Part 2: "The Ethnography of Communication", 133–136.
- Goffman, E. 1974. *Frame Analysis: An Essay on the Organization of Experience*. Cambridge, MA: Harvard University Press.
- Hassabis, D. 2016. *The Future of Artificial Intelligence, Strachey Lecture 2016, Computer Science Series*. Oxford, UK: Oxford University. URL: <https://podcasts.ox.ac.uk/artificial-intelligence-and-future> (accessed November 30, 2022).
- Heller, A. 1990. *Can Modernity Survive?* Berkeley, CA: University of California Press.
- Johnson, D.G., and M. Verdicchio. 2017. "Reframing AI Discourse". *Minds and Machines* 27: 575–590.
- Kajiwar, T. 1979. *The Direction of Play*. Tokyo: The Ishi Press.
- Krieg, G., and G. Kohs. 2017. *AlphaGo. The Movie*. Moxie Pictures. URL: <https://www.youtube.com/watch?v=WXuK6gekU1Y> (accessed November 1, 2022).
- Lampel, J. 2001. "Show-and-Tell: Product Demonstrations and Path Creation of Technological Change". In *Path Dependence and Creation*, edited by R. Garud and P. Karnoe, 303–327. Mahwah: Lawrence Erlbaum.
- Latour, B. 1993. *We Have Never Been Modern*. Cambridge, MA: Harvard University Press.
- Latour, B. 1996. "Social theory and the study of computerized work sites". In *Information Technology and Changes in Organizational Work. IFIP Advances in Information and Communication Technology*, edited by W. J. Orlikowski, G. Walsham, M. R. Jones, and J. I. Degross, 295–307. Boston, MA: Springer. DOI: 10.1007/978-0-387-34872-8\_18.
- LeCun, Y., Y. Bengio, and G. Hinton. 2015. "Deep learning". *Nature* 521(7553): 436–444.
- Lynch, M. 2022. "Garfinkel's Studies of Work". In *The Ethnomethodology Program: Legacies and Prospects*, edited by D. W. Maynard, and J. Heritage, 114–38. Oxford, UK; New York, NY: Oxford University Press.
- Lynch M., E. Livingston, and H. Garfinkel. 1983. "Temporal Order in Laboratory Work". In *Science Observed: Perspectives on the Social Study of Science*, edited by K. Knorr Cetina, and M. Mulkey, 205–238. London, UK: Sage.
- Mair, M., P. Brooker, W. Dutton, and P. Sormani. 2021. "Just what are we doing when we're describing AI? Harvey Sacks, the commentator machine, and the descriptive politics of the new artificial intelligence". *Qualitative Research* 21(3): 341–359.
- Marres, N. 2020. "Co-existence or displacement: Do street trials of intelligent vehicles test society?". *The British journal of sociology* 71(3): 537–555.
- Marres, N., M. Guggenheim, and A. Wilkie, Eds. 2018. *Inventing the Social*. Manchester, UK: Mattering Press.
- Marres, N., and D. Stark. 2020. "Put to the test: For a new sociology of testing". *British Journal of Sociology* 71(3): 423–443.
- Neff, G., and P. Nagy. 2016. "Talking to bots: Symbiotic agency and the case of Tay". *International Journal of Communication* 10: 4915–4931.
- Pollner, M. 2012. "Reflections on Garfinkel and Ethnomethodology's Program". *American Sociologist* 43(1): 36–54.
- Quéré, L. 1998. "The still-neglected situation?". *Réseaux: The French journal of communication, Communication-Technologie-Société* 6(2): 223–253.
- Quéré, L., and C. Terzi. 2015. "Pour une sociologie pragmatiste de l'expérience publique", *SociologieS* [online], published 23 February 2015, consulted 14 February 2023. URL: <http://journals.openedition.org/sociologies/4949>.
- Roberge, J., and M. Castelle. 2021. "Toward an End-to-End Sociology of 21st-Century Machine Learning". In *The Cultural Life of Machine Learning*, edited by J. Roberge, and M. Castelle, 1–29. Cham: Palgrave Macmillan.
- Sacks, H. 1992. *Lectures on Conversation*. Oxford, UK: Basil Blackwell.
- Sharma, S. and P. Brooker. 2016. "#notracist: Exploring racism denial talk on Twitter". In *Digital Sociologies*, edited by J. Daniels, K. Gregory, and T.M. Cottom, 463–485. Bristol, UK: Policy Press.
- Smith, W. 2015. "Technologies of stage magic: Simulation and dissimulation". *Social Studies of Science* 45(3): 319–343.
- Sormani, P. June 2018. "Logic-in-Action? AlphaGo, Surprise Move 37 and Interaction Analysis". In *Handbook of the 6th World Congress and School on Universal Logic*, edited by J.-Y. Beziau, A. Buchsbaum, and C. Rey, 378. Vichy: Université Clermont Auvergne.
- Sormani, P. 2019. "Reenactment as a Research Strategy: From Performance Art to Video Analysis and Back Again". In *Dialogues Between Artistic Research and Science and Technology Studies*, edited by H. Borgdorff, P. Peters, and T. Pinch, Chapter 13. London, UK; New York, NY: Routledge.

- Sormani, P. 2022. "Remaking Intelligence? Of Machines, Media, and Montage". *Tecnoscienza. Italian Journal of Science and Technology Studies* 13(2): 57-86.
- Star, S.L. 1989. "The Structure of Ill-Structured Problems: Boundary Objects and Heterogeneous Distributed Problem Solving". In *Distributed Artificial Intelligence*, edited by L. Gasser, and M. H. Huhns, 37-54. London, UK: Pitman Publishing.
- Star, S. L. 1999. "The ethnography of infrastructure". *American behavioral scientist* 43(3): 377-391.
- Suchman, L. 1987. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge, UK: Cambridge University Press.
- Suchman, L. 2008. "Feminist STS and the Sciences of the Artificial". In *The handbook of science and technology studies*, 3rd. Edition, edited by E. J. Hackett, O. Amsterdamska, M. E. Lynch, and J. Wajcman, 139-164. Cambridge, MA: MIT Press.
- Suchman, L. (2007). *Human-Machine Reconfigurations. Plans and Situated Actions*, 2<sup>nd</sup> Edition. Cambridge, UK: Cambridge University Press.
- Sudnow, D. 1978. *Ways of the Hand: The Organization of Improvised Conduct*. Cambridge, MA: MIT Press.
- Winner, L. 1977. *Autonomous Technology: Technics-Out-of-Control as a Theme in Political Thought*. Cambridge, MA: MIT Press.
- Woolgar, S. 2004. "What Happened to Provocation in Science and Technology Studies?". *History and Technology* 20(4): 339-349.



## Authors

**Noortje Marres** is Professor of Science, Technology and Society at the Centre for Interdisciplinary Methodology at the University of Warwick and Guest Professor in the Media of Cooperation Research Centre at the University of Siegen. She published two monographs *Material Participation* (2012) and *Digital Sociology* (2017). She is PI on the international ORA project *Shaping 21st Century AI: Controversy and Closure in Research, Policy and Media* and received a Leverhulme Fellowship for the project *Beyond the Lab: An empirical philosophy of intelligent vehicle testing in the UK*.

**Philippe Sormani** is Senior Researcher and Co-Director of the Science and Technology Studies Lab at the University of Lausanne. Drawing on and developing ethnomethodology, he has published on experimentation in and across different fields of activity, ranging from experimental physics (in *Re-specifying Lab Ethnography*, 2014) to artistic experiments (in *Practicing Art/Science*, 2019). Currently, he is experimenting with 'DIY AI', educational technology, and media studies.