



From United Steel to Waymo: industrializing simulation

Sam Hind¹

Received: 8 April 2024 / Accepted: 5 August 2024

© The Author(s) 2024

Abstract

The use of computers for simulation work can be traced back to the 1950s, and the pioneering work of Stafford Beer, KD Tocher and others at Cybor House in Sheffield, UK, the research and development (R&D) department of British steelmakers, United Steel. This innovative simulation work sought to offer an abstracted, ‘total’ environment of the steelmaking process in which different operational activities could be modeled. Critical to this work was the ability of computer simulations to perform such modelling at a fraction of the cost, wasting fewer material resources, and in a considerably shorter timeframe. Such work can be understood as the earliest example of the application of industrial-scale ‘automated computation’ to a real-world industrial process. Similarly indebted to the early principles of computer simulation, Waymo engineers are also engaged in the building of so-called ‘conflict typologies’ designed to encode material properties of everyday driving interactions between road users, rather than simply road users themselves. Through ‘motion planning’, coupled with the categorization of driving interactions, Waymo engineers build instrumental understanding of their own system’s purported intelligence in navigating everyday driving situations. Functioning as ‘generative mechanisms’ rather than simply evaluative devices, engineers seek to industrialize—instrumentalize, scale up, rationalize—everyday driving knowledge. Through conflict typologies, instrumental knowledge of the actual capacities of autonomous vehicles is industrialized, materialized, and realized.

Keywords Simulations · Autonomous vehicles · Industrialization · Machine learning · Classification · Typologies

1 Introduction

The use of computers for simulation work can be traced back to the 1950s, and the pioneering work of Stafford Beer, KD Tocher and others at Cybor House in Sheffield, UK, the research and development (R&D) department of British steelmakers, United Steel. This innovative simulation work sought to offer an abstracted, ‘total’ environment of the steelmaking process in which different operational states, activities, and scenarios could be modeled and tested. Critical to this work was the ability of computer simulations to perform such modeling and testing at a fraction of the cost, wasting fewer material resources, and in a considerably shorter timeframe.

Mapping the steelmaking process—from the pouring of steel to the casting of ingots—was an important step in the materialization, and realization, of the value of simulation to

organizational decision-making. In essence, such work can be understood as the earliest example of the application of industrial-scale ‘automated computation’ (Pasquinelli 2023, p. 41) to a real-world industrial process. Key to this was the computer simulations’ ability to offer ‘instrumental understandings’ (Skemp 1978) of everyday workplace activities, while remaining ‘domain-agnostic’ (Ribes et al. 2019).

This ‘classification work’ is integral to contemporary machine learning (Bechmann and Bowker 2019), and the ‘industrialisation of artificial intelligence’ (van der Vlist et al. 2024, p. 1) more broadly. In autonomous driving, simulations offer huge analytical possibilities—generating material evidence of the ‘intelligence’ of their machine learning-dependent autonomous vehicle systems. For Waymo, Google/Alphabet’s autonomous vehicle division, extrapolating ‘what if’ scenarios in simulation software offers engineers the chance to model counterfactuals based on real situations (Schwall et al. 2020). These ‘synthetic’ simulated events consequently produce a welter of different insights into the capabilities of their vehicle systems, each useful for calibrating these systems—bridging the ‘reality

✉ Sam Hind
Sam.hind@manchester.ac.uk

¹ University of Manchester, Manchester, UK

gap’ (Steinhoff and Hind 2024)—to the driving world at large.

Similarly indebted to the early principles of computer simulation, Waymo engineers are also engaged in the building of so-called ‘conflict typologies’, a form of computational capture (Agre 1994) designed to encode material properties of everyday driving interactions between road users, rather than simply road users themselves. Through ‘motion planning’, coupled with the categorization of driving interactions, Waymo engineers build instrumental understanding of their own system’s purported intelligence in navigating everyday driving situations.

By instrumental understanding, I mean operational, rule-based knowledge of a specific technological process—whether steelmaking or autonomous driving—deemed necessary for its successful execution. Instrumental knowledge is the kind of knowledge that swirls around, circulates through, and pools in particular places and settings within an organizational or institutional context (Lave and Wenger (1991). Engineers—whether adept at early computer simulation or contemporary machine learning—are those who are typically engaged in cultivating, standardizing, materializing, and instrumentalizing such knowledge within a workplace setting.

Conflict typologies, far from just retrospective tools for evaluating past crashes—so-called ‘contact events’—instead govern and guide the machine learning work carried out by Waymo engineers to optimize their own vehicles. Through these ‘generative mechanisms’ engineers seek to industrialize—instrumentalize, scale up, rationalize—everyday driving knowledge. Through conflict typologies instrumental knowledge of the actual capacities of autonomous vehicles are industrialized, materialized, and realized.

2 United Steel: simulation work in industrial settings

Simulation work can be understood as work-based activities centered on designing, optimizing, and utilizing computer simulations (Küppers et al. 2006). Often understood through the lens of cybernetics (Wiener 1962; Pickering 2009) and cybernetic management (Beer 1959), computers began to be used for simulation work in industrial settings because of their potential ability to map ‘closed-loop’ factory production activities such as steelmaking. KD Tocher’s *The Art of Simulation* (1963) offered the technical grounding for such early computer simulation work, building on prior work on sampling methods and random number generation.

As Tocher (1963) suggested, computer simulation work was indebted to three intersecting trajectories: the theory of mathematical statistics, applied mathematics involving

partial differential equations, and the ‘new science’ (Tocher 1963, p. 3) of Operational Research (OR). At Cybor House Tocher and colleagues had amassed expertise in all three areas and more.¹ As Hollocks recalls, the staff at Cybor House, on establishment in 1957, ‘included three psychologists, an anthropologist, two zoologists, a philosopher and a classicist—as well as the range of scientific disciplines more normally (now) associated with an Operational Research department’ (Hollocks 2006, p. 19). It was an eclectic mix of scientific disciplines and expertise, reflective of the need for novel ideas, and interdisciplinary thinking, in the application of new computational ideas to specific industrial problems.

Steinhoff and Hind (2024) refer to the early (non-digital computer) era of simulation as the ‘statistical regime’ (Steinhoff and Hind 2024, p. 7), in which pseudo-random number generation in the form of ‘Monte Carlo simulation’ was typically employed. Realizing its value, it quickly shifted from being used purely as a ‘numerical calculation machine’ to an ‘alternate reality...on which “experimentation” could be conducted’ (Galison 1996, p. 119, quoted in Steinhoff and Hind 2024, p. 7). The second era, the ‘discrete-event regime’ (Steinhoff and Hind 2024, p. 8), offered an enhanced environment in which Monte Carlo-style simulations ‘progressively gave way to more involved bespoke models of real systems’ (Hollocks 2008, p. 131, quoted in Steinhoff and Hind 2024, p. 8). Cybor House ‘became the key location where [this] cutting-edge simulation work started to be done’ (Steinhoff and Hind 2024, p. 8), still based on theories of mathematical statistics, where a specific industrial sites’ ‘normal’ operating behavior could be modeled, but where more expansive, and experimental, forms of computer simulation were conducted.

2.1 Productivity and agglomeration

The use of early computer simulations within industrial settings offers a perspective on how, and where, machine learning-driven simulation might be employed now. As the United Steel example suggests, the application of computer simulations stimulated increased economic productivity, enabled by economic agglomeration. In 1957, the year Cybor House opened, British steel production output was 22 million tonnes (Office for National Statistics 2016). In 1971, four years after the British steel industry was nationalized, there were 302,600 steel workers, reaching peak steel production output in 1970 of 28.3 million tonnes

¹ Cybor House was a portmanteau of cybernetics and operational research (CY-BOR).

(Office for National Statistics 2016).² Innovations such as computer simulation became of huge potential value to the industry helping to rationalize production processes, lower production costs, and maximize production outputs.

Akin to an early Silicon Valley, Sheffield was known as ‘Steel City’, home to steelmaking since the eighteenth century. Cybor House was located in the Broomhill area of Sheffield, with United Steel operating steelmaking sites across South Yorkshire and Lincolnshire, such as the famous Appleby Frodingham Works in Scunthorpe, and a site at Templeborough located between Sheffield and Rotherham (Hollocks 2006). Located just over an hour away from Manchester, where the first stored-program computers were being developed (at the University of Manchester) and manufactured (at Ferranti) (Lavington 2019), United Steel took advantage of a proto-economic agglomeration of electrical engineering research, computing manufacturing, and the steel industry concentrated around Manchester and Sheffield, in the north of England (Duranton and Kerr 2018; Klepper 2010; Warren 1969).³

2.2 Domains and mechanization

The application of computer simulation within the steel industry can be understood in two further ways: first, as an early example of the ‘logic of domains’ (Ribes et al. 2019) that now pervades the computing industry. As Ribes et al. understand it, the concept of the domain first became prevalent during AI research on ‘expert systems’ during the 1960s and 1970s, where ‘the concept of the domain serves to objectivize the knowledge of circumscribed groups in order to “capture” and “encode” it within expert systems’ (Ribes et al. 2019, p. 284). As Woolgar suggested, expert systems were ‘computer programmes intended to serve as consultants for decision making’ (Woolgar 1985, p. 560–561), rather than fully ‘autonomous’ agents in any real sense. Accordingly, such expert systems required two parts: a ‘knowledge base containing the facts and heuristics of a particular discipline’ (Woolgar 1985, p. 561) otherwise known as ‘domain knowledge’, and an ‘inference procedure, a set of rules for the manipulation of the knowledge base’ (Woolgar 1985, p. 561) in order for the expert system to be able to assist in domain-specific decision-making.

Pervading this and other meanings of domains is the distinction between ‘domain specificity’ and domain

independence’ (Ribes et al. 2019). As Ribes et al. state, ‘making the crossing from independence to domains or vice versa, is a central concern for those adopting the logic. Domain independent tools, techniques, algorithms or theory must be “applied”, “tailored”, or “customized” to a specific domain’ (Ribes et al 2019, p. 284). The General Steelplant Program (GSP) developed at United Steel was an early demonstration of this logic in action, understood as ‘perhaps...the most significant work of enduring value to be carried out at Cybor House’ (Hollocks 2006, p. 21). Over time, the GSP morphed into a domain independent, General *Simulation* Program for use in other settings beyond steel manufacturing. As Hollocks recounts, United Steel were drawn to using such a simulation program precisely because it offered a ‘general specificity’: the potential to map production cycles across the different furnace methods in use across their multiple plants. As Hollocks writes:

[KD] Tocher knew that United Steel had a number of plants across the north of England. The plants in Scunthorpe, Rotherham, Sheffield and Workington covered three different technologies: open-hearth, electric arc and Bessemer converter...So Tocher saw the challenge as in producing a comprehensive model that could be used for any of these sites—a General Steelplant Program, GSP....Although there was clearly a similarity in purpose across the United Steels’ steel plants, the various technologies, equipment and layouts meant important differences in modelling. Tocher thus had to conceive a framework that would address the steel plant problem more generically. (Hollocks 2008, p. 132)

Second, the development of the GSP, and the general specificity it offered, can be placed within a longer history of the mechanization of both hand calculation and mental labor (Pasquinelli 2023). This mechanization involves, thus, not only an ‘industrialization’ of calculation (i.e., a ‘scaling up’), but the incorporation of specific representational forms—maps, plans, and as such was the case with United Steel, diagrams or programs. As Pasquinelli considers, ‘the idea of the automatic computer, in the contemporary sense, emerged out of the project to mechanise the mental labour of clerks rather than the old alchemic dream of building thinking automata...’ (Pasquinelli 2023, p. 40).

The question for United Steel, for which a combination of a Ferranti Pegasus computer and their in-house GSP were the answer, was how to use new computing technologies to aid decision-making across all operations, using these technological innovations to rationalize steel production (Lavington 2000). Here, what Pasquinelli refers to as ‘automated computation’ (Pasquinelli 2023, p. 41) concerned both an industrialization of computation, and the subsequent application of industrialized computation

² By comparison, 34,500 people were employed in the steel industry in 2014, and steel production output was 11 million tonnes in 2016 (Office for National Statistics 2016). The annual volume of crude steel produced in the UK was 5.96 million tonnes in 2022 (Statista 2022).

³ Many of the immediate steelmaking enterprises lay in the Lower Don Valley area of Sheffield/Rotherham, next to the River Don.

to (heavy) industry. Computer simulation, thus, can be considered an extension of these principles and practices, seeking to offer the computation of possibilities in aid of industrial rationalization. Simulation offered industrial actors, and those acting on their behalf (i.e., managers), the ability to play with metaphorical levers without wasting time, labor, and resources pulling literal equivalents.

2.3 Instructional language

As mentioned in the previous section, computer simulation built on the prior use of statistical methods. As Hollocks wrote, ‘across business and industry, manual/“hand” simulation was a not uncommon tool in OR and Work Study departments’ activity in the 1950s/1960s, using tables of random numbers as the foundation’ (Hollocks 2008, p. 131). The innovation was a computerization, and above all an *industrialization*, of simulation itself. Three aspects were key to this industrialization: the development of an instructional language, the implementation and instrumentation of resulting knowledge, and an imagined, idealized vision of how it might further be put to use.

First, as a form of discrete-event simulation, the GSP considered the steelplant ‘as a set of machines, each with a set of states’ (Tocher 1960, p. 59):

Any change(s) of the state(s) of (a) machine(s) is (are) regarded as an event and the simulation moves from event to event. At any moment of times, machines are grouped together in activities, which endure for a sampled time, and then become free, after a possible change of state, to regroup with other machines in further activities. (Tocher 1960, p. 59)

The development of an instructional language was critical to a kind of ‘instrumental understanding’ (Skemp 1978) of the discrete events integral to the steelmaking process. As Tocher wrote, ‘a language has been developed for naming the machines systematically and describing their states and the times of changes in these. Tests on the states and changes to them can be made in statements in language’ (Tocher 1960, p. 59). In this case, instrumental understanding would amount to knowledge of, and ability to execute, the rules governing steel production in any one instance or setting.⁴ This instructional language was accompanied by a visual flow diagram of all steelplant activities [Fig. 1], constituting a holistic ‘operational ontology’ encompassing all steelplant sites and technologies.

⁴ Throughout, I use instrumental understanding in a slightly different sense to Skemp (1978)—not necessarily as a kind of rule-based knowledge without context or relations, but as a kind of situated rule-based knowledge, dependent on the specifics of setting, in the sense intended by Lave and Wenger (1991).

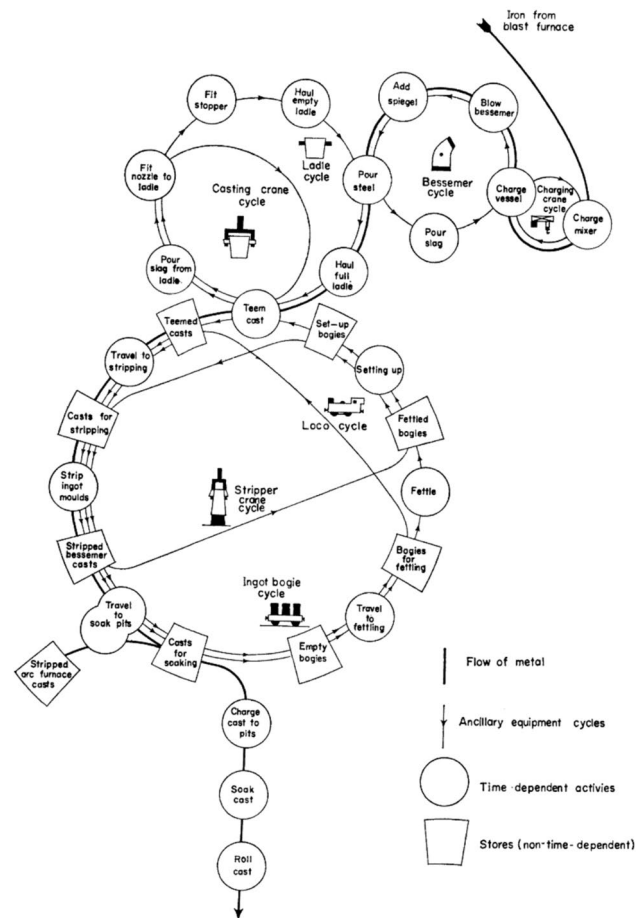


FIG. 4. Simplified flow diagram of activities. Acid Bessemer steel-making plant.

60

Fig. 1 Simplified flow diagram of activities. Acid Bessemer steel-making plant. Source: Tocher (1960)

2.4 Instrumentation of knowledge

Relevant to Tocher’s (1960) articulation of discrete-event simulation is Philip Agre’s (1994) notion of computational capture, as defined during the age of desktop computing in the 1990s. It might typically comprise of five stages: analysis, articulation, imposition, instrumentation, and elaboration. As Agre suggested, the first analytical stage involves the study of ‘an existing form of activity’ identifying ‘its fundamental units in terms of some ontology (entities, relations, functions, primitive actions, and so forth)’ (Agre 1994, p. 109–110). In the next stage, a ‘grammar’ of the activity is established, a way ‘in which [the] units can be strung together to form actual sensible stretches of activity’ (Agre 1994, p. 110). As Agre contends, establishing this ‘grammar of action’ (Agre 1994, p. 109) is far from straightforward, ‘often requir[ing] revision of the

preceding ontological analysis' (Agre 1994, p. 110). Once figured out, 'the resulting grammar is then given a normative force' (Agre 1994, p. 110), with those workers engaged in the activity 'induced to organize their actions so that they are readily "parsable" in terms of the grammar' (Agre 1994, p. 110). The two final subsequent stages—instrumentation and elaboration—concern the ongoing execution and maintenance of the computational capture process, with data generated through it 'stored, inspected, audited, merged with other records, [and] subjected to statistical analysis' (Agre 1994, p. 110). Agre's illustration of how existing activities (i.e., within an office setting) might be captured and rendered through an operational ontology, was already being considered by early computer simulation practitioners like Tocher.

Agre's (1994) positing of computational capture can further be understood as an update to the 'Babbage principle', established by Charles Babbage in 1832, that

states that the organisation of a production process into small tasks (the division of labour) allows for the calculation and precise purchase of the quantity of labour that is necessary for each task (the division of value). The division of labour establishes a privileged perspective for the surveillance of labour, but also helps to modulate the extraction of surplus labour from each worker according to need. In more analytical terms, the Babbage principle posits that the abstract diagram of the division of labour helps to organise production while at the same time offering an *instrument* for measuring the value of labour. In this respect, the division of labour provides not only the design of machinery but also of the business plan. (Pasquinelli 2023, p. 47, authors' emphasis)

In an empirical illustration of Agre (1994) and the Babbage principle, Tocher (1960) outlined eight stages for the use and embedding of the GSP in specific work contexts. These can be divided into three corresponding phases that map onto Agre's (1994) own capture model. In the first analytical/articulation phase, plant processes were studied, and a simulation model mapping those processes is designed. Next, in an imposition phase, a decision control procedure was developed, plant records were collected and analyzed, and simulated models and control procedures were translated to a computer program. Lastly, once deployed, in an instrumentation/elaboration phase, long-term operational goals were evaluated, an experimental 'goal searching' programme might be designed, and equipment for 'real' plant control might be manufactured (Tocher 1960, p. 64–65).

Two points can be made on the above. First, that multiple different workers were involved in the instrumentation of knowledge. Earlier stages were conducted by engineers and

'staff trained in simulation model building' (Tocher 1960, p. 64), with 'work study teams' and programmers engaged in the extraction of plant record data (Tocher 1960, p. 64). While 'plant managers' (Tocher 1960, p. 65) dealt with long-term goals, 'laboratory staff', plant managers and 'equipment manufacturers' (Tocher 1960, p. 65) ideally worked together to design new plant equipment.

Second, that the actual stages involved were necessarily more complex, and more specific, than Agre's (1994) model supposes—by its very nature. In the case of the GSP, it was clear that the 'normative force' of the model Agre (1994, p. 110) speaks of, concerned not only 'real-time capture' (Agre 1994, p. 110) but also a need for 'security, efficiency, protection from liability, and simple control' (Agre 1994, p. 110) as he likewise mentions. Here the value of the GSP was in its 'bespoke generality'—being able to be deployed in different contexts, following an established, executable blueprint.

2.5 Idealized vision of elaboration

Some of Tocher's (1960) model was evidently speculative, an imagined, idealized vision of how the GSP might further be used. In relation to the final stage, Tocher wrote that it had 'not yet been reached in any applications of this technique' (Tocher 1960, p. 65). Nevertheless, he made a proposal, in which specific machines under both human and automatic observation could be connected with data fed into a computer to be outputted as punched tape data, where a 'special computer programme provides for the checking and initial sorting of the data...[which] can be followed by any desired standard programmes for data analysis' (Tocher 1960, p. 64). With this, Tocher (1960) envisioned a 'real-time' application of computer simulation, and how it could be integrated into new computational—evidently cybernetic—systems of real-time data collection, activity analysis, and control. In short, that simulation would play a significant part in the design and implementation of industrial forms of the kind of self-regulating, feedback-enabling, cybernetic systems imagined by Beer (1959).

3 Industrialization of AI

As a bridge to the next section, it is important to understand how the industrialization of computing has shaped computer simulation and AI/machine learning alike. As Steinhoff and Hind have written, contemporary synthetic data work, indebted to earlier innovations in computer simulation, is itself 'part of a long history of the industrialization of computing technology' (Steinhoff and Hind 2024, p. 16). Likewise, that proliferating dependency on cloud computing services offered by the likes of Microsoft Azure, Google

Cloud Platform (GCP), and Amazon Web Services (AWS) constitutes the wholesale ‘industrialisation of artificial intelligence’ (van der Vlist et al. 2024, p. 1). The question of computation—whether in the form of computer simulation, synthetic data, or machine learning—is never too far away from the question of industrialization.

How, then, can the industrialization of such technologies be characterized? Firstly, this involves a ‘scaling up’ of particular technologies from experimental or test environments to real-world settings (Pfothenauer et al. 2022). Contemporary big tech firms such as Microsoft, Google/Alphabet and Amazon are similarly referred to as ‘hyper-scalers’ (Narayan 2022), able to use aforementioned cloud computing resources to offer planetary-scale services to clients. In this, the ability to scale up particular technologies is usually framed as a kind of entrepreneurial zeal or innovative ambition, long the preserve of modernist projects, but supercharged ‘in the era of big tech, [where] the aim is frequently to scale up first and profit later’ (Pfothenauer et al. 2022, p. 5).

Second, this requires a certain commitment to, and actualization of, capital investment—both in the technology itself and the expertise to develop, maintain, and adapt it. Contemporary generative AI products demand huge resources to train underlying large-language models (LLMs) and to deploy them in specific settings. OpenAI’s GPT-4 cost over US\$100million to train, according to its CEO Sam Altman (Knight 2023), with future models likely to cost considerably more (Murgia and Hammond 2024), depending on the number of parameters they utilize (Patel 2023). To connect to the previous issue, ‘scaling out these models to users and agents costs far too much’ (Patel and Wong 2023, n.p.), with the expense of feeding live domain data into machine learning models (known as ‘inference’) ‘exceed[ing] that of training by multiple folds’ (Patel and Wong 2023, n.p.).

Third, it demands a ready—or at least plausible—activity to which it can be applied. For LLMs, the activities are various, with OpenAI’s GPT-3 evaluated on a range of text-based tasks, from next-word prediction and ‘closed book’ question answering, to arithmetic and news article generation (Brown et al. 2020). GPT-4 already underpins Microsoft’s Copilot, an AI-driven assistant designed to help users with an array of rudimentary computer-based office tasks, from creating agendas and drafting boilerplate emails, to summarizing documents and creating presentations (Microsoft 2024).

3.1 Classification work in machine learning

Classification work is integral to the industrialization of AI. As Muldoon et al. write, ‘AI data workers are required for a variety of different tasks in the AI production process, from

very early stages of data collection and organisation, up to the final stages of model evaluation and data verification’ (Muldoon et al. 2024, p. 8). AI data work, as Muldoon et al. (2024) term it, might well be considered a form of ‘microwork’ (Irani 2015), ‘cloudwork’ (Woodcock and Graham 2019) or even ‘ghost work’ (Gray and Suri 2019), although this is certainly not always the case. Indeed, classification work in machine learning typically straddles these definitions, depending on the context. In the case study to follow, I argue that the design of ‘conflict typologies’ in autonomous driving simulations is a form of classification work, echoing the development of operational ontologies in industrial settings in the 1950s. Despite being integral to machine learning, as Tubaro et al. (2020) consider, it can hardly be considered a form of microwork if this involves tasks that are ‘barely visible and poorly compensated’ (Tubaro et al. 2020, p. 2) or involve ‘activities that humans can do quickly and easily’ (Tubaro et al. 2020, p. 2).

Bechmann and Bowker (2019) consider classification work as integral to the building of machine learning models—a critical stage in the AI ‘pipeline’. Different machine learning approaches operationalize classification in different ways, with ‘supervised’ approaches concerning ‘algorithms that work with classifiers or labels to generate predefined outputs’ (Bechmann and Bowker 2019, p. 4) and ‘unsupervised’ approaches with ‘no predefined output... no such supervisor... [and] only the input data’ (Alpaydin 2016, p. 5). The objective of this classification work is to determine a certain ‘ground truth’ (Jaton 2021) on which a machine learning model can be built, dividing the world into constituent pieces and objects, each with their own qualities and characteristics. As a process of ‘localization’ (Gil-Fournier and Parikka 2021), establishing a ground truth for a machine learning model anchors its interpretive abilities to a statistical real-world, much like previous iterations of computer simulation (Steinhoff and Hind 2024). In the absence of classification work, and specifically, ‘without the training dataset as a ground truth, there is no way in which a specific ML [machine learning] model or method can be judged as accurate or, indeed, successful...’ (Hind 2024, p.76).

Classification work is not only found in contemporary machine learning settings. As Bowker and Leigh Star suggest, ‘to classify is human’ (Bowker and Leigh Star 1999, p. 1), and despite the thoroughly modernist practice of classification work, ‘not all classifications take formal shape or are standardized in commercial and bureaucratic products’ (Bowker and Leigh Star 1999, p. 1). In our day to day lives, ‘we have certain knowledge of...intimate spaces’ (Bowker and Leigh Star 1999, p. 2) where such everyday categorization work takes place: ‘any part of the home, school, or workplace reveals...systems of classification’ (Bowker and Leigh Star 1999, p. 2). Accordingly,

classifications ‘appear to live partly in our hands—definitely not just in the head or in any formal algorithm’ (Bowker and Leigh Star 1999, p. 2). Knowledge, and understanding, of the categories and category choices we make are enrolled into specific practices, ‘embodied in a flow of mundane tasks...and many varied social roles’ (Bowker and Leigh Star 1999, p. 2).

Despite the ordinary invisibility of what Bowker and Leigh Star call a kind of ‘folk classification’ (Bowker and Star 1999, p. 2), ‘the formal, bureaucratic ones trail behind them the entourage of permits, forms, numerals, and the sometimes-visible work of people who adjust them to make organizations run smoothly’ (Bowker and Star 1999, p. 2). In such settings, classifications are more evident, more noticeable—attached to formal documentation, defined in job roles, work policies, and as illustrated before, laid down in simulation program manuals. Those who perform classification work might not, themselves, be classified quite so plainly as ‘classification workers’, but much of this work—especially forms of office work and ‘tech’ work—comprises a significant part of what they do. The question then becomes: ‘but what *are* these categories? Who makes them, and who may change them? When and why do they become visible? How do they spread?’ (Bowker and Star 1999, p. 3, authors’ emphasis). As they further suggest:

No one, including Foucault (1970, 1972), has systematically tackled the question of how these properties inform social and moral order via the new technological and electronic infrastructures. Few have looked at the creation and maintenance of complex classifications as a kind of work practice with its attendant financial, skill, and moral dimensions. (Bowker and Leigh Star 1999, p. 5)

This is despite the moral and cultural forces that not only power them, but bake them into ‘the modern information technology world’ (Bowker and Leigh Star 1999, p. 5). As Bechmann and Bowker contend, ‘categories are not a priori constructed, but highly context sensitive’ (Bechmann and Bowker 2019, p. 4), reflecting the specific social, cultural, political, and organizational settings of their creation and maintenance. As contemporary examples of ‘new technological and electronic infrastructures’, machine learning is replete with classifications. While classification work does not occur equally across different machine learning approaches, as mentioned earlier, as Bechmann and Bowker contend, ‘we are still talking about classification work all the way down—the only issue is how visible and how a priori that work is’ (Bechmann and Bowker 2019, p. 4). Accounting ‘for how classes and social categorization arise in the design process as deliberate [as well as] unintentional consequences of decisions made’ (Bechmann and Bowker 2019, p. 4) is critical. Classification work

does not simply disappear in the absence of established or communicated categories: it bubbles under the surface or goes by another name.

Machine learners ‘are often simply called “classifiers”’ (Mackenzie 2017, p. 10) because of the importance of classification work to the machine learning model building process. While machine learning is dependent upon prior classification work, it also generates ‘new categorical workings or mechanisms of differentiation’ (Mackenzie 2017, p. 10). The ‘learning’ part of machine learning, thus, is largely an assumed ability to ‘invent or find new sets of categories for...particular purpose[s]’ (Mackenzie 2017, p. 10). If deemed useful in some way, these new categories are put to work themselves, dividing different social activities and interactions up based on their exhibited qualities. This classification work—how it is performed, who it is carried out by—can be understood to be a product of the industrialization of AI and machine learning, with its incessant, and infinite, need to classify objects and phenomena. Thus, while the existence and utility of cloud computing infrastructures are the most evident illustrations of the industrialisation of AI (van der Vlist et al. 2024), evidence can also be found in the form, scale, and type of classification work being conducted in particular settings.

As Mackenzie considers, paying ‘attention to [the] specificity of practices is an elementary prerequisite to understanding human–machine relations and their transformations’ (Mackenzie 2017, p. 9). Thus, as Mackenzie further suggests:

If we understand machine learning as a data practice that reconfigures local centers of power and knowledge by redrawing human-machine relations, then differences associated with machine learners in the production of knowledge should be a focus of attention. (Mackenzie 2017, p. 9-10)

To reiterate, as Bechmann and Bowker (2019) consider, classification work is integral to the building of machine learning models. Here, data collection, data cleaning, and model training are all classificatory steps in the machine learning process, where machine learners might simply ‘assume’ or derive a priori categories from ‘institutionalized or accepted knowledges’ (Mackenzie 2017, p. 10).

3.2 Classification work in the automotive industry

As a bridge to the next section, it is important to provide an insight into how classification work is performed in the automotive industry. As Tubaro and Casilli contend, the automotive industry ‘has become one of the largest clients of digital data-related micro-working services, notably for the development of autonomous vehicles and of connected cars’ (Tubaro and Casilli 2019, p. 335). This

microwork might take different forms within the industry, but typically involves different kinds of classification work. As Tubaro and Casilli consider, specific classification tasks might involve image classification (organizing images according to certain criteria like day/night), object detection or tagging (labeling road users like cars and cyclists), landmark detection (identifying features in the wider driving environment) or semantic segmentation (raster-based, pixel-level categorization) (Tubaro and Casilli 2019, p. 340). In the aggregate, these tasks ultimately constitute a ground truth on which machine learning models depend, dividing a real-world driving environment into constituent parts, upon which a subsequent autonomous vehicle system can calculate the ‘path trajectories’ of other vehicles (Hind 2023).

Yet, what is important to consider here is that while such work is crucial to machine learning, and necessary to build autonomous vehicle systems, much of it within the automotive industry is not necessarily outsourced as forms of crowd or cloudwork. Instead, such work more typically is conducted in universities and research centers, occupying a fuzzy space between manufacturers and public institutions (Hind et al. 2024). Most of the AI microwork firms identified by Tubaro and Casilli are understood to be ‘generalist’ platforms (Tubaro and Casilli 2019, p. 337), i.e., that they do not offer classification services specifically, or only, for automotive clients. The now-shuttered Argo AI Center for Autonomous Vehicle Research, hosted at Carnegie Mellon University (CMU) in Pittsburgh, USA demonstrated this distinction, being home to research on image classification and ‘sensor work’ (Hind 2023) more broadly. Other autonomous vehicle start-ups such as Waabi occupy a similar space, led by Raquel Urtasun, a computer scientist based at the University of Toronto.⁵ In such cases, fine-grain image annotation work is typically required for autonomous vehicle applications, more typically carried out by computer science graduates in research settings, such as with development of the foundational KITTI Vision Benchmark Suite.⁶ The setting of machine learning and machine vision ‘challenges’ follows a similar blueprint, enrolling newly-minted PhD students in tackling machine learning problems in the autonomous driving domain (Hind et al. 2024). For all these reasons, such classification work is not outsourced, and not conducted by remote workers in the sense articulated by the likes of Tubaro et al. (2020) and Muldoon et al. (2024). This is despite new forms of automotive microwork emerging, such as remote ‘intervention’ by engineers (Hind 2022a) to help rescue stranded autonomous vehicles (Hawkins 2023).

4 Waymo: Waymo driver and contact events

In this final section, I contend that the design of ‘conflict typologies’ for autonomous driving simulations are intended to standardize, scale, and ultimately industrialize AI within this specific domain context. In February 2023, engineers in Google/Alphabet’s autonomous vehicle division released a research paper examining the safety performance of the Waymo Driver, Waymo’s autonomous vehicle system (Victor et al. 2023). In an accompanying blogpost, Waymo announced they had ‘accomplished another first’ exceeding ‘one million miles on public roads with no human behind the wheel’ (Waymo 2023, n.p.). Both documents can be understood as key components in Waymo’s efforts to demonstrate the safety of their autonomous vehicles, in a bid to convince the wider public, government, legislators, and regulators that their vehicles are fit for public use.

















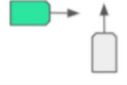
More specifically, they can be understood as evidence of Waymo’s work to both ‘suspend’ and ‘manage’ the meaning of crashes—referred to as ‘contact events’ by the company—involving their vehicles, in order to provide a convincing illustration of their safety (Hind 2024). This and other public releases of safety data from their vehicles (Schwall et al. 2020) can be considered responses to past events involving autonomous vehicles, like the death of Elaine Herzberg—hit by an Uber ATG autonomous vehicle undergoing testing—in 2018 (Hind 2022b; Smiley 2022). In late 2023, ‘robotaxi’ operator Cruise had its license suspended by the California Department of Motor Vehicles, after obstructing an investigation into a crash involving one of its autonomous vehicles (Korosec 2023). In short, ‘winning’ the discursive battle over the purported safety of autonomous vehicles has been an increasingly important aim for autonomous vehicle firms.

The purpose of the research paper, as stated by the authors, was firstly ‘to examine all contact events experienced during the first one-million miles of rider-only (RO) operations of the Waymo Driver’ (Victor et al. 2023, p. 7) and second to ‘explore what conclusions can be made from this observed real-world safety performance in terms of the frequency and severity of these contact events’ (Victor et al. 2023, p. 7). Waymo conducts autonomous vehicle operations in two locations in the US (Phoenix, Arizona and San Francisco) using two versions of their Waymo Driver platform, referred to as the fourth generation (Chrysler Pacifica) and fifth generation (Jaguar I-Pace). Both versions are active in Phoenix, with only the fifth generation deployed in San Francisco. Analyzing data collected from all one-million miles of RO operations, from 2019 to January 21, 2023, the authors identified 20 contact events. Of these, 1 occurred in 2020, 6 in 2021, 11 in 2022, and 2 in 2023. 18 occurred in Phoenix, with another two in San Francisco. 12

⁵ <https://waabi.ai/>.

⁶ <https://www.cvlibs.net/datasets/kitti/>.

Fig. 2 Eight of the 16 conflict groups and short descriptions. Source: Victor et al. (2023)

Conflict Group	Picture	Description
Single Vehicle (SV)  		Includes all actions (or lack thereof) where the ego vehicle is traveling in a trafficway but then experiences an in-trafficway interaction without a conflict partner (e.g., a rollover event) or an off-trafficway interaction (e.g., a road departure).
Front-to-Rear (F2R) 		Involves one road user interacting with another road user in the same direction and same travel lane.
Same-Direction Lateral Incursion (SDLI) 		Occurs when two roadway actors are traveling in the same trafficway but in initially different travel lanes at the time of the initial interaction due to lateral incursion by some actor.
Same-Direction Prior Circumstances (SDPC) 		Involves two roadway actors operating on the same trafficway in the same direction when one road user performs a lateral evasive action, experiences loss of control, or is involved in a prior collision that results in an interaction with the other road user.
Opposite Direction Lateral Incursion (ODLI) 		Occurs when a non-turning actor operating in the trafficway's intended travel direction interacts with another actor that is operating opposite of the travel direction in the same trafficway.
Opposite Direction Prior Circumstances (ODPC) 		Involves two roadway actors traveling in opposite direction trafficways in their respective trafficway's direction of travel when one road user performs a lateral evasive action, experiences loss of control, or is involved in a prior collision that results in an interaction with the other road user.
Turn into Path Opposite Direction (TIPOD) 		Occurs as a result of one actor changing vehicle-operated trafficways via a turning maneuver and interacting with another actor, where one of these actors is operating in the opposite direction of the trafficway's direction of travel.
Intersection Cross Traffic (ICT) 		Involves interactions that occur as a result of both actors changing or crossing over trafficways, and where the two actors cross paths with one another.

contact events involved fourth generation versions of Waymo Driver (i.e., Chrysler Pacifica), and a further eight involved fifth generation versions (i.e., Jaguar I-Pace) (Victor et al. 2023).

All recorded contact events are collated into a single table, summarizing much of the above details, and including a 'danger description' and a 'narrative description' of each contact event (Victor et al. 2023, p. 11–13). A calculation of injury risk accompanies each entry, using an adapted version of the industry-standard Maximum Abbreviated Injury Scale of 2 or greater (MAIS2+), to determine the probability of an AIS2 level injury or higher ($p(\text{MAIS2+})$). Examples of these injuries include 'concussions with no or brief loss of conscience, fractures to the sternum, and 2 or fewer rib fractures' (Victor et al. 2023, p. 8). The first contact event registered a ($p(\text{MAIS2+})$) score of 4%, eight contact events are calculated between 1 and 2%, and a final 11 contact events register a score between 0 and 1% (Victor et al. 2023, p. 11–13).

4.1 Conflict typologies as industrializing knowledge

Each contact event is categorized into one of 16 'conflict groups' which comprises an overarching 'conflict typology' [Fig. 2]. As the authors state, 'the conflict groups are one of the layers of a conflict typology that also describes the conflict partners, role (initiator or responder), and the perspectives of each actor involved in a conflict' (Victor et al. 2023, p. 28). In developing this broad conflict landscape, 'a conflict typology can be used in safety impact methodologies that analyze and predict the potential performance of a safety countermeasure or system within a set of defined crash modes' (Kusano et al. 2023, p. 1).⁷

The value of the conflict typology is in how it is used to categorize contact events. Each conflict group demarcates a different contact event. The most familiar, or easy to

⁷ Kusano et al. also mention that conflict typologies are sometimes called 'crash groups', 'maneuver typologies' or 'scenario typologies' (Kusano et al. 2023, p. 1).

imagine, of these might be single-vehicle (SV) conflict, described as including ‘all actions (or lack thereof) where the ego vehicle is traveling in a trafficway but then experiences an in-trafficway interaction without a conflict partner (e.g., a rollover event) or an off-trafficway interaction (e.g., a road departure)’ (Victor et al. 2023, p. 28). Five of the 20 reported incidents are categorized as SV events. In addition, a Front-to-Rear (F2R) conflict is described as involving ‘one road user interacting with another road user in the same direction and same travel lane’ (Victor et al. 2023, p. 28). Of the contact events reported, a total of six are classified as F2R events. Other entries are categorized as Backing (BACK) events, denoting ‘all interactions where at least one road user is moving in reverse’ (Victor et al. 2023, p. 29). Eight incidents are classified as BACK events. The single remaining documented contact event falls into the Opposite Direction Lateral Incursion (ODLI) category, described as occurring ‘when a non-turning actor operating in the trafficway’s intended travel direction interacts with another actor that is operating opposite of the travel direction in the same trafficway’ (Victor et al. 2023, p. 29).

The narrative description of this final event, occurring in Phoenix in 2022 and involving a fifth-generation vehicle (i.e., Jaguar I-Pace), offers an insight into its categorization:

Contact occurred between the left rear corner of a Waymo AV [autonomous vehicle] and the side of a garbage truck. The Waymo AV had pulled to the right on a narrow residential street, unable to proceed past an upcoming garbage truck and a parked vehicle. While attempting to pass the Waymo AV, the garbage truck made contact with the left rear corner of the Waymo AV. At the time of contact, the Waymo AV was stationary and the garbage truck was traveling less than 1 mph. (Victor et al. 2023, p. 12)

4.1.1 Abstraction and discretization

Waymo engineers have designed a far more extensive categorization of conflicts than those actually recorded by their own vehicles over one-million RO miles. The 20 contact events recorded only fall into one of four categories: SV, F2R, BACK and ODLI. A further 12 categories are not represented in the analyzed contact event data, and according to the typology created, no other conflicts are recorded at all. While there may be various reasons for this, what appears evident is that the conflict typology is designed to be utilized more broadly than to simply classify actually occurring contact events.

As a follow-up to earlier work where the documentation of simulated contact events was detailed (Schwall et al. 2020), the typology can be understood as a *generative mechanism* rather than merely a ‘safety impact evaluation’

tool (Kusano et al. 2023, p. 3) for assessing, and classifying, prior contact events. In other words, it governs and guides machine learning work being performed by Waymo engineers, as they optimize their autonomous vehicle systems in question.

The construction of a conflict typology can be understood as an intended totalizing abstraction of all possible vehicle interactions—in the same way that the GSP could be understood as a technique for mapping all steelmaking activities. While the nature of the activities being captured and categorized is markedly different—steelmaking as a ‘closed’ system, vehicle interactions as an ‘open’ system—the design and documentation of a conflict typology hints at the underlying operational desire to formulate driving, and contact events, as a series of *discrete* events. This is what they refer to in related work as ‘scenario-based testing’ (Kusano and Victor 2022), i.e., a methodology ‘to estimate the probability of an injury outcome in a scenario’ (Kusano and Victor 2022, p. S225) in which an autonomous vehicle is involved. While the typology arguably ‘naturally evolves as novel scenarios are encountered’ (Kusano et al. 2023, p. 3), it nonetheless operates as a holistic tool.

While the exact way in which conflict typologies are used in simulation work at Waymo is not divulged, Kusano et al. write that conflict typologies might be used alongside ‘scenario description languages’ (Kusano et al. 2023, p. 20) which ‘focus on describing scenarios in a way that can be translated into simulations or evaluations of an ADS [automated driving system]’ (Kusano et al. 2023, p. 20). More specifically, that:

The conflict typology could be used in conjunction with a set of scenarios to organize them by actor types, groups, perspectives, and contributing factors. For example, this conflict typology is the basis for the aggregation used in the Collision Avoidance Testing scenario-based testing programme at Waymo, where collision avoidance competency is evaluated relative to a reference behaviour model in conflicts where the ADS is the responder role vehicle. (Kusano et al. 2023, p. 20-21).

In other words, that an instructional language (i.e., conflict groups) might be used to implement future autonomous vehicle simulations.

4.1.2 Interactions

Second, and necessarily then, conflict typologies can be considered as operating ‘downstream’ from ground truth-oriented classification work. Following the metaphor of the AI pipeline, commonly used within the domain of autonomous driving, this conflict typology work is typical of the tasks being conducted during a motion planning/forecasting

phase, where engineers work to optimize an autonomous vehicle system's ability to forecast the future trajectories of other vehicles and road users (Hind 2023). In such cases, there is a clear interest in properly attributing, and classifying certain characteristics not only of specific vehicle types (i.e., a car, truck, bus) but also the types of *vehicle interactions* they might be implicated in, susceptible to, or commonly involved in. In other words, that a garbage truck—like in the ODLI example above—might be predisposed to certain kinds of conflicts, based on its size, weight, poor manoeuvrability, and the activities it is typically engaged in (i.e., collecting garbage). As a result, the autonomous vehicle system being tested might require further fine-tuning in order to account for the increased likelihood and risk of being involved in these incidents—whether the fault of the autonomous vehicle or not. Waymo's Collision Avoidance Testing (CAT) is where such work might specifically be carried out, with Waymo testing 'all scenarios used for the safety evaluation of [their] latest software releases in simulation—whether derived from test track data, real-world data, or synthetic means' (Waymo 2022, n.p.). Here, conflict typologies feed into scenario-based testing, conducted through computer simulations.

4.1.3 Patterns

In designing and applying a conflict typology, knowledge about recorded contact events is both materialized and abstracted. Lacking the typology, and in the absence of work to categorize events into specific conflict groups, the events stand as singular artifacts. Sorted into classes, they constitute prospective *patterns* to be joined by subsequent incidents, categorized accordingly. In building these classes with more entries, Waymo engineers slowly come to understand their vehicle's capacities and limitations differently: as those potentially prone to involvement in certain kinds of conflicts over others. The categorization of the 20 contact events into 4 conflict groups (SV, F2R, BACK, ODLI) suggest that Waymo vehicles are prone to single-vehicle, front-to-rear, backing, and opposite direction lateral incursion events—whether, ultimately, they are blame or not. Mapping these capacities appears key to instrumental understandings of autonomous vehicles, critical to simulating—and avoiding—future conflicts.

To summarize, not only are such objects being classified—common enough to image classification work across many different contexts—but that object *interactions* are being classified in addition. Here, objects (road users) themselves might be categorized in ground truth training datasets as possessing certain geometric properties (length, width), but this ground truth-based classification work is—in these documented conflict typologies—being accompanied with downstream vehicle interaction classification work.

Both are critical, but the latter seems especially novel in this context, seeking to further discretize other aspects of driving.

5 Conclusion

In this article, I have considered how forms of 'classification work' in the development of autonomous vehicle systems industrialize knowledge. Through the categorization of 'contact events' involving their autonomous vehicles, Waymo is able to operationalize understanding of what kinds of interactions their vehicles are involved in. Rather than a post-facto evaluative framework, 'conflict typologies' are key to the advance modeling of autonomous vehicle interactions: they govern and guide the work of Waymo engineers.

Classification work is everywhere in machine learning. Without the labeling and annotation of training data, machine learning models cannot be built (Engdahl 2024). To gain real-world applicability, they rely on the hard labor of data annotators—sometimes crowdsourced, sometimes remote, sometimes piecemeal. Yet this classification work can take many forms. In this instance, the building of conflict typologies can be considered as somewhat different, conducted in-house. Such classification work requires close attention to the final application, and knowledge—I argue here—of the applied domain in question. In such instances, knowledge of how these conflict typologies are to be used in the simulation of autonomous vehicle systems is critical. The 'counterfactual calculations' detailed by one Waymo team (Schwall et al. 2020) support the conflict typology work detailed by another (Victor et al. 2023)—and vice versa.

This contemporary work—classification, typologies, machine learning—seemingly novel, is 'part of a long history of the industrialization of computing technology' (Steinhoff and Hind 2024, p. 16). The current industrialization of AI (van der Vlist et al. 2024), dependent upon the 'hyper-scaling' of cloud computing capacities (Narayan 2022), can be situated within a 70-year history of the industrialization of digital computing. Rather than a later addition, computer simulation has always figured within this history. The industrialization of computing technology, then, has always involved the industrialization of computer simulation. What is striking is that the industrialization of computer simulation thus offered opportunities for *industrial application*, such as in the dominant UK steel industry during the 1950s. While strange to consider from a contemporary vantage point, the steel industry was the perfect setting for innovative computer simulation work.

United Steel, a steelmakers based in Sheffield, UK, pioneered the application of computer simulations to

the modeling of industrial processes like steelmaking. In Cybor House, the company's innovative R&D department, Stafford Beer, KD Tocher and colleagues developed a General Steelplant Program (GSP), offering the possibility of mapping the steelmaking process across three different sites in the north of England (Templeborough, Scunthorpe, Stocksbridge), based on three different steelmaking technologies (open hearth, electric, arc, acid Bessemer). In simulating production processes at a level of abstraction capable of accounting for each site and technology, while still remaining useful for making production decisions, the GSP constituted something even more consequential: a General *Simulation* Program (Steinhoff and Hind 2024). Able to move between domain-specificity and domain-agnosticism, this new program offered the tangible prospect of being able to use computer simulation to rationalize production.

In mapping the flows, activities, and states of the steelmaking processes, GSP operators developed an 'operational ontology', key to computer simulation's role in cultivating, standardizing, materializing, and instrumentalizing any activity it was instructed to model. This computational capture process, as Agre (1994) later documents in relation to office-based work settings, exercised the Babbage principle, dividing the labor process up into non-divisible units of activity, in order to determine and extract productive value (Pasquinelli 2023).

Articulating these historic connections between computing, simulation, industrial application, and classification work is important in order to contextualize contemporary machine learning practices. Understanding where instrumental knowledge is generated, ordered, shared, and implemented in different work contexts is key to understanding how different computational technologies—including machine learning and AI—have long-sought to classify and categorize phenomena in order to control and manage it. AI's industrial roots in the automation of mental labor, as Pasquinelli (2023) examines, is part of its social history. As the history of the industrialization of computing, and the industrial application of computer simulation, suggest, this classification and categorization work is not distributed equally among all participants. Whether 'staff trained in simulation model building' (Tocher 1960, p. 64) in the north of England in the late 1950s, or software engineers in Mountain View, California in 2024, understanding the capacity and ability for certain people to perform classification work, requires specific, empirical analysis. In these cases and all others, those who perform such tasks—the mapping of steelplant activities, the creation of conflict typologies—do so as employees of specific kinds of firms. In this, computer simulation has always been a tool for managerial decision-making, deployed to rationalize production in any way possible (Hind 2024).

Consequently, three fruitful avenues of future research are worth mentioning. First, more work needs to be done on the industrial history of *computer simulation*—where it was deployed, what type of rationalization work it performed, and what kinds of managerial decisions it supported. While the work at United Steel in the 1950s and 60s was innovative at the time, developments pushed them aside. Indeed, computer simulation work moved into new industries and sectors altogether, with new visual-interactive approaches offering unprecedented advances in mobility settings (airports, train networks) in the 1980s (Steinhoff and Hind 2024). These sectoral shifts remain underexamined, yet help to complicate understanding of how the trajectory of computer simulation unfolded, and incorporated new decision-making theories and frameworks.

Second, analysis of the ensuing *industrialization of AI* is evidently necessary, from the scaling-up of cloud computing (Narayan 2022; van der Vlist et al. 2024), and the foundational role of machine learning benchmark datasets (Engdahl 2024), to the platformization of AI challenges (Luitse et al. 2024; Hind et al. 2024). This industrialization of AI demands the cultivation, standardization, materialization and instrumentalization of new kinds of operational knowledges which ultimately benefit the 'rentier' strategies of big tech firms (Christophers 2020).

Lastly, deepening scrutiny of the politics and ethics of *synthetic data* and generative AI is needed (Steinhoff 2022; Jacobsen 2023; Helm et al. 2024). While typically seen as a kind of 'fix', either avoiding the collection of sensitive user data or lowering the cost of collecting training data, synthetic data generation must still wrestle with bridging the so-called 'reality gap' between real world and simulation. The statistical rules that are meant to govern '*sim-to-real* transfer' (Salvato et al. 2021) rely on establishing the norms of distribution in any given context—from 'realistic' driving scenarios (Wayve 2023) to 'representative' medical image data (Guo et al. 2024). These so-called 'foundation models'—rather than eradicating these messy issues—will only complicate the political and ethical effects of computer simulation as they embed themselves in various sites and settings in the future.

Acknowledgements The author would like to thank the special issue co-editors for organizing the panel, 'Where's the "Intelligence" in AI? Mattering, Placing and De-individuating AI' at STS Italia 2023, Bologna (Italy), and for their subsequent invitation to submit an article for this special issue.

Funding This work was supported by the German Research Foundation (*Deutsche Forschungsgemeinschaft, DFG*), project number 262513311 (*SFB 1187: 'Media of Cooperation'*).

Data availability Not applicable.

Declarations

Conflict of interest The corresponding author(s) state that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agre P (1994) Surveillance and capture: two models of privacy. *Inf Soc* 10(2):101–127. <https://doi.org/10.1080/01972243.1994.9960162>
- Alpaydin E (2016) *Machine learning*. MIT Press, Cambridge
- Bechmann A, Bowker GC (2019) Unsupervised by any other name: hidden layers of knowledge production in artificial intelligence on social media. *Big Data Soc* 6(1):1–11. <https://doi.org/10.1177/2053951718819569>
- Beer S (1959) *Cybernetics and management*. The English Universities Press, London
- Bowker G, Leigh Star S (1999) *Sorting things out: classification and its consequences*. MIT Press, Cambridge
- Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, Neelakantan A, Shyam P, Sastry G, Askell A, Agarwal S, Herbert-Voss A, Krueger G, Henighan T, Child R, Ramesh A, Ziegler DM, Wu J, Winter C, Hesse C, Chen M, Sigler E, Litwin M, Gray S, Chess B, Clark J, Berner C, McCandlish S, Radford A, Sutskever I, Amodei D (2020) Language models are few-shot learners. *ArXiv*. <https://doi.org/10.48550/arXiv.2005.14165>. Accessed 27 March 2024
- Christophers B (2020) *Rentier capitalism: who owns the economy, and who pays for it?* Verso, London
- Duranton G, Kerr W (2018) The logic of agglomeration. In: Clark GL, Feldman M, Gertler MS, Wójcik D (eds) *The New Oxford Handbook of Economic Geography*. Oxford University Press, Oxford, pp 347–365
- Engdahl I (2024) Agreements ‘in the wild’: standards and alignment in machine learning benchmark dataset construction. *Big Data Soc* 11(2):1–13. <https://doi.org/10.1177/20539517241242457>
- Foucault M (1970) *The order of things: an archaeology of the human sciences*. Tavistock Publications, London
- Foucault M (1972) *The archaeology of knowledge*. Tavistock Publications, London
- Galison P (1996) Computer simulations and the trading zone. In: Galison P, Stump DJ (eds) *The disunity of science: boundaries, contexts, and power*. Stanford University Press, Stanford, pp 118–157
- Gil-Fournier A, Parikka J (2021) Ground truth to fake geographies: machine vision and learning in visual practices. *AI Soc* 36:1253–1262. <https://doi.org/10.1007/s00146-020-01062-3>
- Gray ML, Suri S (2019) *Ghost work: how to stop silicon valley from building a new global underclass*. Houghton Mifflin Harcourt, Boston
- Guo P, Yang D, Zhao C, Xu D (2024) Addressing medical imaging limitations with synthetic data generation. *Nvidia Developer*. <https://developer.nvidia.com/blog/addressing-medical-imaging-limitations-with-synthetic-data-generation/>
- Hawkins AJ (2023) Robotaxi companies have a serious trust issue. *The Verge*. <https://www.theverge.com/23948708/cruise-robotaxi-suspension-trust-remote-assist>. Accessed 29 March 2024
- Helm P, Lipp B, Pujadas R (2024) Generating reality and silencing debate: synthetic data as discursive device. *Big Data Soc* 11(2):1–5. <https://doi.org/10.1177/20539517241249447>
- Hind S (2022a) Making decisions: the normal interventions of Nissan ‘mobility managers.’ *Mobilities* 17(4):467–483. <https://doi.org/10.1080/17450101.2021.1988682>
- Hind S (2022b) Machinic sensemaking in the streets: More-than-lidar in autonomous vehicles. In: Rose G (ed) *Seeing the city digitally: reconfiguring temporalities, spatialities and collectivities*. Amsterdam University Press, Amsterdam, pp 57–80
- Hind S (2023) Sensor work: enabling the interoperation of autonomous vehicles. *J Comput-Mediat Commun* 28(5):1–10. <https://doi.org/10.1093/jcmc/zmad014>
- Hind S (2024) *Driving decisions: how autonomous vehicles make sense of the world*. Palgrave Macmillan, London
- Hind S, van der Vlist FN, Kanderske M (2024) Challenges as catalysts: how Waymo’s Open Dataset Challenges shape AI development. *AI & Society* 1–17. <https://doi.org/10.1007/s00146-024-01927-x>
- Hollocks BW (2006) A glimpse of early OR—A brief history of Cybor House. *Or Insight* 19(1):18–25. <https://doi.org/10.1057/ori.2006.4>
- Hollocks BW (2008) Intelligence, innovation and integrity – KD Tocher and the dawn of simulation. *J Simulat* 2(3):128–137. <https://doi.org/10.1057/jos.2008.15>
- Irani L (2015) The cultural work of microwork. *New Media Soc* 17(5):720–739. <https://doi.org/10.1177/1461444813511926>
- Jacobsen BN (2023) Machine learning and the politics of synthetic data. *Big Data Soc* 10(1):1–12. <https://doi.org/10.1177/20539517221145372>
- Jaton F (2021) *The Constitution of Algorithms: Ground-Truthing, Programming, Formulating*. MIT Press, Cambridge
- Klepper S (2010) The origin and growth of industry clusters: the making of silicon valley and detroit. *J Urban Econ* 67(1):15–32. <https://doi.org/10.1016/j.jue.2009.09.004>
- Knight W (2023) OpenAI’s CEO says the age of giant AI models is already over. *Wired*. <https://www.wired.com/story/openai-ceo-sam-altman-the-age-of-giant-ai-models-is-already-over/>. Accessed 27 March 2024
- Korosec K (2023). California DMV immediately suspends Cruise’s robotaxi permit. *TechCrunch*. <https://techcrunch.com/2023/10/24/dmv-immediately-suspends-cruises-robotaxi-permit-in-california>. Accessed 27 March 2024
- Küppers G, Lenhard J, Shinn T (2006) Computer simulation: Practice, epistemology, and social dynamics. In: Lenhard J, Küppers G, Shinn T (eds) *Simulation: pragmatic construction of reality*. Springer, Dordrecht, pp 3–22
- Kusano K, Scanlon JM, Brännström M, Engström J, Victor T (2023) Framework for a conflict typology including contributing factors for use in ADS safety evaluation. *Proceedings of the 27th International Technical Conference on the Enhanced Safety of Vehicles (ESV)* 1–23. <https://www.esv.nhtsa.dot.gov/Proceedings/27/27ESV-000328.pdf>
- Kusano K, Victor T (2022) Methodology for determining maximum injury potential for automated driving system evaluation. *Traffic Inj Prev* 23(Sup1):S224–S227. <https://doi.org/10.1080/15389588.2022.2125231>
- Lave J, Wenger E (1991) *Situated learning: legitimate peripheral participation*. Cambridge University Press, Cambridge
- Lavington S (2000) *The pegasus story: a history of a vintage british computer*. Science Museum, London
- Lavington S (2019) *Early Computing in Britain: Ferranti Ltd. And Government Funding, 1948-1958*. Springer Cham, Cham

- Luitse D, Blanke T, Poell T (2024) AI competitions as infrastructures of power in medical imaging. *Inform Commun Soc*. <https://doi.org/10.1080/1369118X.2024.2334393>
- Mackenzie A (2017) *Machine learners: archaeology of a data practice*. MIT Press, Cambridge
- Microsoft (2024) Copilot lab. Microsoft. <https://copilot.cloud.microsoft/en-us/prompts>. Accessed 27 March 2024
- Muldoon J, Cant C, Wu B, Graham M (2024) A typology of artificial intelligence data work. *Big Data Soc* 11(1):1–13. <https://doi.org/10.1177/20539517241232632>
- Murgia M, Hammond G (2024) Can OpenAI create superintelligence before it runs out of cash? *Financial Times*. <https://www.ft.com/content/6314d78d-81f3-43f5-9daf-b10f3ff9e24f>. Accessed 27 March 2024
- Narayan D (2022) Platform capitalism and cloud infrastructure: theorizing a hyper-scalable computing regime. *Environ Plan A* 54(5):911–929. <https://doi.org/10.1177/0308518X221094028>
- Office for National Statistics (2016) Updated: The British steel industry since the 1970s. Census 2021. <https://www.ons.gov.uk/economy/economicoutputandproductivity/output/articles/updatedthebritishsteelindustrysincethe1970s/2016-01-18>. Accessed 14 March 2024
- Pasquinelli M (2023) *The eye of the master: a social history of artificial intelligence*. Verso, London
- Patel D, Wong G (2023) GPT-4 architecture, infrastructure, training dataset, costs, vision, MoE. *SemiAnalysis*. <https://www.semianalysis.com/p/gpt-4-architecture-infrastructure>. Accessed 27 March 2024
- Patel D (2023) The AI brick wall: A practical limit for scaling dense transformer models, and how GPT 4 will break past it. *SemiAnalysis*. <https://www.semianalysis.com/p/the-ai-brick-wall-a-practical-limit>. Accessed 27 March 2024
- Pfotenhauer S, Laurent B, Papageorgiou K, Stilgoe J (2022) The politics of scaling. *Soc Stud Sci* 52(1):3–34. <https://doi.org/10.1177/03063127211048945>
- Pickering A (2009) *The cybernetic brain: sketches of another future*. The University of Chicago Press, Chicago
- Ribes D, Hoffman AS, Slota SC, Bowker GC (2019) The logic of domains. *Soc Stud Sci* 49(3):281–309. <https://doi.org/10.1177/0306312719849709>
- Salvato E, Fenu G, Medvet E, Pellegrino FA (2021) Crossing the reality gap: a survey on sim-to-real transferability of robot controllers in reinforcement learning. *IEEE Access* 9:153171–153187. <https://doi.org/10.1109/ACCESS.2021.3126658>
- Schwall M, Daniel T, Victor T, Favarò F, Hohnhold H (2020) Waymo public road safety performance data. *ArXiv*. <https://doi.org/10.48550/arXiv.2011.00038>. Accessed 28 March 2024
- Skemp RR (1978) Relational understanding and instrumental understanding. *The Arithmetic Teacher* 26(3): 9–15. <https://www.jstor.org/stable/41187667>
- Smiley L (2022) ‘I’m the operator’: The aftermath of a self-driving tragedy. *Wired*. <https://www.wired.com/story/uber-self-driving-car-fatal-crash/>. Accessed 28 March 2024
- Statista (2022) Annual volume of crude steel produced in the United Kingdom (UK) from 2009 to 2022. Statista. <https://www.statista.com/statistics/474809/annual-production-of-crude-steel-uk/>. Accessed 14 March 2024
- Steinhoff J (2022) Toward a political economy of synthetic data: A data-intensive capitalism that is not a surveillance capitalism? *New Media Soc* 26(6):3290–3306. <https://doi.org/10.1177/1461448221099217>
- Steinhoff J, Hind S (2024) Simulation and the reality gap: Moments in a prehistory of synthetic data. *MediArXiv*. <https://doi.org/10.33767/osf.io/np3vb>
- Tocher KD (1960) An integrated project for the design and appraisal of mechanized decision-making control systems. *Or* 11(1–2):50–65. <https://doi.org/10.2307/3007180>
- Tocher KD (1963) *The art of simulation*. The English Universities Press, London
- Tubaro P, Casilli AA (2019) Micro-work, artificial intelligence and the automotive industry. *J Ind Bus Econ* 46:333–345. <https://doi.org/10.1007/s40812-019-00121-1>
- Tubaro P, Casilli AA, Colville M (2020) The trainer, the verifier, the imitator: three ways in which human platform workers support artificial intelligence. *Big Data Soc* 7(1):1–12. <https://doi.org/10.1177/2053951720919776>
- Van der Vlist FN, Helmond A, Ferrari FL (2024) Big AI: Cloud infrastructure dependence and the industrialisation of artificial intelligence. *Big Data Soc* 11(1):1–16. <https://doi.org/10.1177/20539517241232630>
- Victor T, Kusano K, Gode T, Chen R, Schwall M (2023) Safety performance of the Waymo Rider-Only Automated Driving System at one million miles. Waymo. <https://storage.googleapis.com/waymo-uploads/files/documents/safety/Safety%20Performance%20of%20Waymo%20RO%20at%201M%20miles.pdf>. Accessed 28 March 2024
- Warren K (1969) Recent changes in the geographical location of the British steel industry. *Geogr J* 135(3):343–361. <https://doi.org/10.2307/1797323>
- Waymo (2022) Waymo’s Collision Avoidance Testing: Evaluating our driver’s ability to avoid crashes compared to humans. Waymo. <https://waymo.com/blog/2022/12/waymos-collision-avoidance-testing/>. Accessed 3 April 2024
- Waymo (2023) First million rider-only miles: How the Waymo Driver is improving road safety. Waymo. <https://waymo.com/blog/2023/02/first-million-rider-only-miles-how.html>. Accessed 28 March 2024
- Wayve (2023) Introducing GAIA-1: A cutting-edge generative AI model for autonomy. Wayve. <https://wayve.ai/thinking/introducing-gaia1>. Accessed 23 July 2024
- Wiener N (1962) [1948] *Cybernetics: or control and communication in the animal and the machine*. MIT Press, Cambridge
- Woodcock J, Graham M (2019) *The gig economy: a critical introduction*. Polity, Cambridge
- Woolgar S (1985) Why not a sociology of machines? The case of sociology and artificial intelligence. *Sociology* 19(4):557–572. <https://doi.org/10.1177/0038038585019004005>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.