Information, Communication & Society

ISSN: 1369-118X (Print) 1468-4462 (Online) Journal homepage: www.tandfonline.com/journals/rics20

Driving the hype: LLMs as ‘general-purpose’ promise in the autonomous vehicle industry

Alex Gekker & Sam Hind

INFORMATION, COMMUNICATION & SOCIETY
https://doi.org/10.1080/1369118X.2026.2689026

Driving the hype: LLMs as 'general-purpose' promise in the autonomous vehicle industry

Alex Gekker [a] and Sam Hind [b]

[a]Department of Media Studies, University of Amsterdam, Amsterdam, The Netherlands; [b]University of Manchester, Manchester, UK

ABSTRACT
Proponents of autonomous vehicles are proclaiming a new era, driven by innovations derived from large language models (LLMs). Following a decade of highly public failures and the concurrent redirection of capital and attention towards generative AI, firms in the autonomous vehicle industry have turned to LLM-style techniques as the basis for a rebooted commercialisation effort. Through a technographic analysis of an annual industry event, Nvidia GTC 2024, and 41 sources of related grey literature (firm communications, technical preprints, trade journalism, and regulatory documents), we examine how two autonomous vehicle firms, Wayve and Waabi, discursively frame these techniques as capable of overcoming the technical, economic, financial, and regulatory limits that have so far thwarted a fully-autonomous future (AV 1.0). We find that these claims draw heavily on the supposed 'general-purposivity' of LLMs, importing assumptions about the broad applicability of foundation models, multimodal training, parallel tokenisation, and epistemic distillation (AV 2.0) into a domain where they remain largely unproven. The innovations discussed are highly speculative in nature and contribute to the discursive hype around the viability of LLM-style approaches in the autonomous vehicle industry. In promoting claims of general-purposivity, they serve as a blueprint for how proponents are seeking to embed LLMs in new industries and domains.

ARTICLE HISTORY
Received 18 July 2025
Accepted 9 June 2026

KEYWORDS
LLMs; general purpose technology; Innovation; Autonomous vehicles; Hype

Introduction

Until large language models (LLMs) took the crown with OpenAI's launch of ChatGPT in November 2022, autonomous vehicles (AVs) were considered the most promising application of artificial intelligence (AI) in the modern era. Huge investments were built on the same bet: that a world of fully-autonomous cars would be possible by 2020 (Metz & Griffith, 2020). Ground-breaking work on robotics that formed the basis of early advances in autonomous vehicles in the early 2000s gave way to a period of

CONTACT Alex Gekker a.gekker@uva.nl Department of Media Studies, University of Amsterdam, PO box 94550, 1090 GN Amsterdam, The Netherlands

'benchmarking' work in the 2010s, ending most recently with the feverish launch of autonomous vehicle start-ups, public tests, and commercial partnerships into the 2020s, designed to deliver autonomous vehicles to the masses (Hind, 2024a). However, a series of highly public failures, coinciding with the concomitant rise of LLMs, has pushed this dream further into the distance.

To demonstrate: in October 2022, Ford shuttered their autonomous vehicle division, Argo AI, absorbing the project into other company initiatives (Hawkins, 2022). In October 2023, California's Department of Motor Vehicles suspended Cruise's 'robotaxi' operating licence in San Francisco, before the company's CEO resigned (Korosec & Bellan, 2023). Then, in February 2024, Apple cancelled its decade-long 'Titan' project, reportedly shifting the workforce to generative AI (Roth, 2024). Together, these incidents constitute a crisis of viability, possibly heralding a dreaded 'AI winter' once prophesied by industry figures (Shalev-Shwartz et al., 2018).

In the time since these announcements, excitement has returned: what if LLMs could breathe life into the stalled revolution? Recent scholarship on the supposed generative AI revolution has emphasised the role of industry hype not only in promoting the technology to consumers but also as essential to the commercialisation process itself (De Kloet, 2026; Kotliar, 2025). This article adds to this growing body of work on the discursive framing of AI, by focusing on the autonomous vehicle industry – a multifaceted collection of firms, consisting of old and new players, that aim to develop commercially viable self-driving cars. Specifically, we examine the claims that LLM-style innovations will lead the autonomous vehicle industry out of AV 1.0 – the first phase of autonomous vehicle commercialisation – and into a brave new era, AV 2.0, due to the technology's general-purpose applicability that counteracts the resources previously required. Put simply: 'what truly transpires in the 'revolutions' and 'disruptions' allegedly brought upon by AI? And what role does hype play in these revolutions?' (Kotliar, 2025, p. 828).

The article proceeds as follows: in section one, we present four innovations – foundations, multimodality, parallel tokenization, and epistemic distillation – that underpin claims to the 'general-purposivity' of LLMs, drawing on work in media studies and cognate fields. Subsequently, we present our 'technographic' analysis (Bucher, 2018; van der Vlist et al., 2024; Woolgar, 1999) of autonomous vehicle grey literature drawn from a study of an annual industry event run by chip firm Nvidia.

In section two, we detail four limits that have thwarted the dream of a fully autonomous future. Firstly, *technical limits* in machine vision and software development. Secondly, *economic limits* in underlying business models and product design. Thirdly, *financial limits* in macroeconomic policies and venture capital investments. And lastly, *regulatory limits* regarding testing incidents and licencing protocols.

In section three, we present two recent examples of new LLM innovations claimed by their developers to be able to overcome some, if not all, of these stated limits. In these cases, innovations from autonomous vehicle start-ups Wayve and Waabi, such as *video prompting* and *parallel tokenization*, are seen as technological breakthroughs that are able to arrest the challenges faced by the autonomous vehicle industry since the rise of generative AI in 2022 – constituting a tempting, if precarious, realisation of the promise of LLMs.

In the discussion and concluding sections, we argue that critical media scholars should be alive to contested claims of the general-purposivity of LLMs and pay close attention to how such innovations are being discursively framed and hyped in specific domains and industries.

Theoretical framework: 'General-purpose' LLM innovations

Since 2021, visual diffusion models such as DALL-E and Stable Diffusion, and subsequent transformer-based text models such as ChatGPT, have sent shockwaves through the AI community and wider world. A key assumption so far has been that AI training requires substantial forms of human labour; for instance, in the curation of training datasets manually annotated by workers, typically in the form of low-paid, piecemeal work conducted remotely in the global south (Muldoon et al., 2024; Tubaro et al., 2020).

Instead, new techniques in unsupervised or self-supervised learning have allowed for the ingestion of massive amounts of unlabelled data and the subsequent creation of robust 'foundation' or general-purpose models capable of generating new outputs based on prompting – a set of instructions often given through a natural language chat interface (Burkhardt & Rieder, 2024). Subsequent shifts into so-called 'agentic' capabilities have further cemented the models' abilities to interact with existing software products, allowing human interlocutors to instruct them in performing tasks beyond text or image generation, mostly through software development (Sapkota et al., 2025).

A key differentiating aspect of the current generation of AI technologies is their supposed 'generality'. Unlike previous iterations of AI and machine learning techniques, pretrained on known datasets to perform specific actions, the transformer and attention architecture underpinning LLMs has, according to proponents, made it applicable to a wide variety of tasks (Luitse & Denkena, 2021). Even more enthusiastically, AI has been heralded as the next General Purpose Technology (Stackpole, 2024), similar to electricity or the Internet (Trajtenberg, 2018). Goldfarb et al. (2023) define a General-Purpose Technology as the confluence of its potential for innovation (subsequent technologies dependent on it) and actual widespread application (amount of people actually using the technology), with an emphasis on specific industries. Based on an empirical analysis of online job data, Goldfarb et al. (2023) argue that machine learning may already be a 'general-purpose' technology because of its evident potential for 'large-scale economic impact' (Goldfarb et al., 2023, p. 1). Similar claims have been made about both generative AI (Steinhoff & Hind, 2025) and LLMs (Eloundou et al., 2023).

The economic promise of generality, combined with soaring valuations and investment in general-purpose model companies such as OpenAI or Anthropic, has led to significant reorganisation of AI research, a redirection of AI funding, and a resetting of AI priorities. Across this period, some have proclaimed an imminent 'existential risk' from AI (Hern, 2024), centred on the prospect of Artificial General Intelligence (AGI) – models capable of surpassing human cognitive abilities. Such claims remain suspect, either through overstatement of perceived 'intelligence' or through the obvious, huge, computational costs and resource limits necessary to 'scale' such AI systems (Narayan, 2022). Some industry commentators now warn that the tide is shifting due to the unsustainable requirements to train the next generation of models compared to their utility (Wong, 2024). Researchers in computer science also caution against grand proclamations, particularly when promised scientific breakthroughs replace existing methods (Narayanan & Kapoor, 2025). Nonetheless, 'regardless of its veracity, hype is an important object of study, and it should be considered in any interrogation of AI' (Kotliar, 2025, p. 16).

In the autonomous vehicle industry, this new hype has challenged the tradition of supervised learning and data annotation. A central shift occurred from relying on mostly human-annotated existent data, to training generalisable models on synthetic, broad data, a process that Mitrokhov (2024) calls a transition from *world models* (that represent the known world) to *model worlds* (that simulate a simplified version of it for agentic training). A new generation of autonomous vehicle firms have emerged that have sought to re-incorporate these LLM innovations, as we discuss later in the article.

In the remainder of this section, we introduce four innovations that we argue are key to the claims being made around the wider applicability and generality of LLMs.

Foundations

Generative AI is based on transformer architectures and *foundation models,* a term that 'captures the "unfinished character" (CRFM, 2022, p. 6) of these systems, and "highlights their role as base infrastructures for various uses and applications"' (Burkhardt & Rieder, 2024, p. 3). Preceding generations of models use neural nets to extrapolate from pre-annotated data or use generative adversarial networks (GANs) and diffusion models to produce realistic images by processing huge amounts of similar pictures. Foundation models have historically not, therefore, been trained for *task-specific* purposes, unlike many preceding approaches. The 'prompting' component typical of foundation models highlights this– requiring a user to specify the task at hand by designing a bespoke query for a foundation model to execute. Whilst this does not mean that foundation models are not trained for any tasks at all – OpenAI's technical report for new models evaluates its performance based on multiple benchmarks, from bar exams to language translation (OpenAI, 2025a)– using a model requires a reasonably skilled design of prompt queries that can extract actual task-specific value from the model itself. However, recent developments suggest greater specification, with firms like OpenAI maintaining parallel model families differentiated by function (OpenAI, 2025b; 2025c) and Apple designing on-device foundation models for task-specific functions rather than general-purpose conversation (Apple, 2025).

Multimodality

Two key components for both training and output of the latest generation of foundation models are their multimodal training and scaling laws. Unlike models trained on text or images only, newer versions are incorporating various types of data in their training corpora (Li et al., 2023), scaling their abilities to work with originally unplanned use cases. Moreover, training new models benefits greatly (on a log scale) from scaling laws, where the more data and parameters one uses, the better the outcome (Kaplan et al., 2020). Initial evidence suggests that multimodal training furthermore improves performance in unexpected ways (Le, 2023). For instance, 'video scraping' innovations have enabled agents like Google's Gemini 1.5 to extract and collate relevant information from screen recordings (e.g., email interfaces) into text (Edwards, 2024), and agents such as OpenAI's Operator and Anthropic's Claude 'computer use' tool can now execute user prompts by autonomously navigating web interfaces or, more recently, controlling a user's entire desktop environment (Anthropic, 2026; OpenAI, 2025c). In such cases, video and interface control are clearly being positioned as 'the next text' by industry actors, despite difficulties in verifying the capabilities of models deployed in limited contexts or under experimental conditions, and acknowledged needs for human oversight (Ferreira, 2026). These approaches have serious ramifications for who can participate in the new multi-modal model space, increasing reliance on Big Tech firms already dominating the generative AI ecosystem (Narayan, 2022; van der Vlist et al., 2024).

Parallel tokenization

When training on a large dataset, developers first use a 'tokenizer' to reduce the corpus to the smallest viable units of meaning. While it's possible to tokenise each single character (or pixel), in practice, this is not feasible for current hardware and software limitations. For text, these are usually (parts of) words that repeat in the corpus. 'Tokenization' thus may be converted to 'token' and 'ization' tokens. Each token then gets transformed into a numerical vector in a multidimensional space (a process called embedding) that encodes their relationship in the corpus. It is through this process that 'meaning' is achieved within the model (Wolf et al., 2020). While the conversion of complex phenomena into discrete bounded tasks is not a new phenomenon in computer sciences, common to both digitalisation (Agre, 1994) and simulation processes (Hind, 2024b; Steinhoff & Hind, 2025) in particular, the distinction here is that tokens are processed in parallel using the attention mechanism rather than sequentially or linearly. This furthers the abstraction level of the model and black-boxes it even from its developers, arguably paving the way for the 'tokenization of everything' by chip firms like Nvidia, as recently reported (Foley, 2025).

Epistemic distillation

A final controversial LLM 'innovation' is its relation to truth. Foundation models tend to 'hallucinate' or 'confabulate' answers, presenting things that are fundamentally untrue, convincingly (Förster, 2023). Explanations for this vary, from the basic principles of LLMs' next-word prediction approach (Douglas Heaven, 2024) to model 'overfitting' and poorly implemented human evaluation 'alignment' processes (Ji et al., 2025). By default, generative models are good at creating 'averages' of their tokenised inputs, 'mean images' that 'represent the norm by signalling the mean' (Steyerl, 2023, n.p.). In practice, models often demonstrate a tangential relationship to the truth, gleaned through the tokenization and embedding process and refined through post-training weighting, rather than in any clear ruleset. Hicks et al. (2024) suggest treating this underlying proposition as 'bullshit' (Frankfurt, 2005) with models being 'agnostic' on truth. Instead, model outputs should only be understood as reflecting correlation within its training data and thus primarily 'portray[ing] itself as a "normal" interlocutor like ourselves' (Hicks et al., 2024, p. 38). Frontier AI labs now routinely invest massive resources in 'mechanistic interpretability', attempting to trace internal epistemic pathways of their models (Ameisen et al., 2025), advocating techniques like 'distillation', where 'student' models learn from 'teacher' models capable of 'transfer[ing] their rich world understanding' (Waymo, 2025, n.p.) to everyday situations. Yet even these efforts have revealed troubling findings: models can fabricate fictitious reasoning processes, and the techniques themselves may work less well for the newest generation of 'reasoning' models that are fast becoming the industry standard (Douglas Heaven, 2026).

Method

Our two-step 'technographic' analysis (Bucher, 2018; van der Vlist et al., 2024; Woolgar, 1999), draws on the study of an annual industry event, Nvidia GTC 2024, referred to as the 'Woodstock of AI' (IN-2025). Held annually in San Jose, California (USA) by the chip firm Nvidia, CEO Jensen Huang typically uses the event to announce new products and innovations, such as advances made in the semiconductor chips used by autonomous vehicle firms. These 'rock concert-style' events have long been used within the tech industry to stimulate excitement and anticipation around new product launches, from Bill Gates and Steve Ballmer's eccentric dancing at the release of Windows 95 to Steve Jobs' iconic announcement of the Apple iPhone at MacWorld 2007.

We follow an established tradition within media studies and science and technology studies (STS) of utilizing industry events as a focal point for analysis of sociotechnical imaginaries (Bucher, 2017; Jasanoff & Kim, 2013) and technological 'hype' (Funk, 2019; Kotliar, 2025), used to examine fields as varied as gambling machines (Schüll, 2014), web cartography (Dalton, 2015), and VR technologies (Egliston & Carter, 2022). We treat industry material not as neutral evidence of technological capability but as discursive interventions that construct particular narratives of innovation, failure, and reinvention, consistent with broader STS and media studies approaches to technological hype and sociotechnical imaginaries.

At Nvidia GTC 2024, a session was hosted by Nvidia Drive, Nvidia's automotive arm, to demonstrate how new AI approaches could benefit – and arguably, were already benefiting – autonomous vehicle development (NV-2024a). In two separate presentations Alex Kendall (CEO of Wayve) (WAY-2024) and Raquel Urtasun (CEO of Waabi) (WAA-2024b) proudly claimed the beginning of a new era of autonomous driving: AV 2.0. In section three we detail both innovations launched at the event: Wayve's GAIA 1.0 and Waabi's CoPilot4D.

Our empirical material comprises 41 sources spanning 2022–2026, drawn from four categories of grey literature: firm communications (product announcements, technical blogposts, investor materials), technical preprints and papers (primarily hosted on ArXiv),[1] trade and broadsheet journalism, and political/regulatory documents. A full list of these sources, classified by type, is provided in Appendix 1. This material was not assembled as a bounded corpus subject to systematic sampling; rather, following established technographic practice, it was gathered iteratively through the study of the two self-proclaimed 'AV 2.0' firms that presented at its Drive session,[2] while investigating the claims made through the sources mentioned above.

These firms were selected not as representative cases of the autonomous vehicle industry at large, but as emblematic proponents of LLM-driven approaches whose public claims around these alleged technical breakthroughs warrant further scrutiny. The four-limit typology (technical, economic, financial, regulatory) emerged through sustained interpretive engagement with this material, rather than through a pre-determined coding scheme. We understand and frame these innovations as examples of how LLM-style approaches are being used by actors in the autonomous vehicle industry to supposedly overcome these limits. In the below, we detail these four limits before evidencing the two applications of LLM innovations in the following section.

The limits to AV 1.0

Technical limits

Technological innovations in machine vision, software development, and fleet management have so far not been able to deliver a fully-autonomous driving experience.

Common rule-based, or 'taxonomical' approaches to machine vision have proved too labour-intensive. These have required the categorisation of every object in a driving environment (bicycle, pedestrian, traffic cone etc.) into discrete 'classes', combined with rules for how autonomous vehicles should react to a growing collection of events, including complex 'sub-rules' for sub-classes of objects (i.e., small and large traffic cones). These approaches have relied on dividing the driving world up into constituent, identifiable parts, based on the translation of sensor inputs (e.g., from cameras) into vector-based outputs (e.g., bounding boxes). These methods have not been able to spot 'occluded' objects (e.g., a cyclist hidden behind a van), requiring additional computational techniques (Figure 1).

Dominant engineering approaches have favoured sequential, modular, development cycles. This has prevented the development of 'interoperable' systems capable of synthesising different data streams to make decisions. Waabi CEO Raquel Urtasun has summarised these issues as related to 'slow development cycles' dependent on 'real-world testing miles' in which vehicle systems still fail to 'generalize skills to new geographies and unseen cases' (WAA-2024b). Prospective 'robotaxi' operators have failed to manage

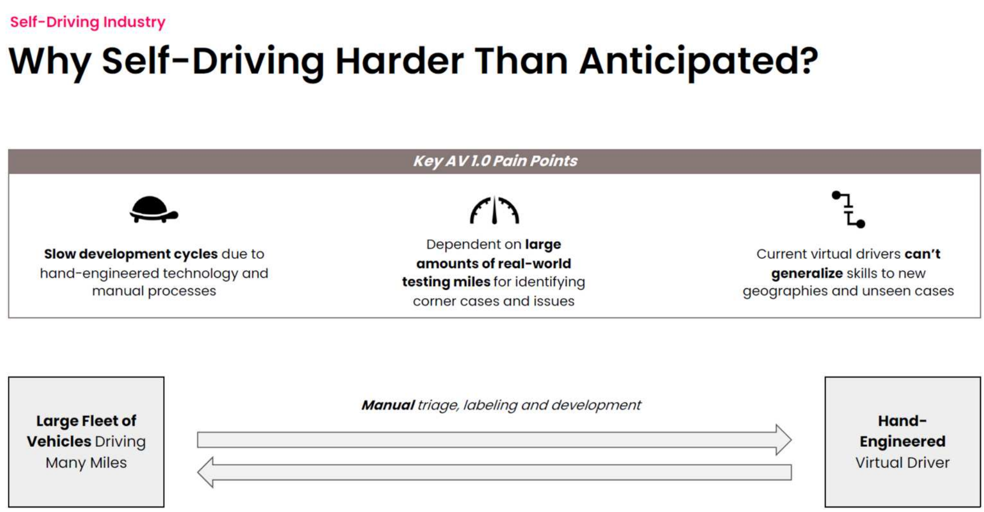

Figure 1. 'Why self-driving harder than anticipated?'. A slide from Raquel Urtasun's presentation at NVIDIA GTC 24 (WAA-2024b).

vehicle fleets. This has delayed the wider public rollout of robotaxis beyond specific test cities such as San Francisco (VE-2023), as operators have underestimated the extent to which vehicle fleets must be maintained, including the regularity with which vehicles break down, and the need for both remote and in-person assistance.

Economic limits

Economic developments in business models, product design, and market competition have also stalled the deployment of autonomous vehicles.

Leading automotive firms such as Toyota, Volkswagen, Honda, Ford, and Hyundai still constitute 29.5% of global automotive sales in 2024 (ST-2025). Such 'legacy' manufacturers have struggled to adapt to new market realities (FI-2024a; WI-2026), opting instead to build driver-assistance products that stop short of offering holistic autonomous capabilities. Ford's BlueCruise system, designed for 'hands-free' motorway driving, can only be activated in designated zones and is made available in select vehicle models (FO-2024a).

At the same time, industry 'disruptors' have so far failed to integrate autonomous vehicles into platform business models. Uber's robotaxi operation (Uber ATG) was acquired by a rival (Aurora) in 2020 after a fatal accident involving one of its vehicles, and has since resorted to partnering with other autonomous vehicle firms (Pony.ai, Waymo) and fleet management operators (Verne) to attempt to scale their robotaxi services across different markets (UB-2023; UB-2026). Chinese electric vehicle firms meanwhile have provided global competition to Western manufacturers. Shenzhen-based BYD has full 'vertical' ownership of their automotive supply chain, from lithium mines to semiconductor production (BY-2022), outsold Tesla in 2023 (FI-2024b), and now constitutes 4.5% of global automotive sales, with a higher market share than both Honda (4.4%) and Ford (4.3%) (ST-2025), prompting the US and EU to levy substantial tariffs on imported Chinese EVs (CN-2024; EU-2024; IE-2025).[3]

Financial limits

Financial developments in macroeconomic policies, venture capital investments, and asset classes have similarly presented challenges to the development of autonomous vehicles.

Global inflationary pressures have dampened tech investments. From 2008 to 2022, tech investments 'flowed freely' (WE-2024) thanks to zero interest-rate policies. Yet since 2022, governments around the world have raised interest rates to control inflation (GU-2023) with the rise of AI and LLMs threatening to reduce remaining available capital to autonomous vehicle projects, agreements, and collaboration (GU-2023).

Automotive manufacturers have attempted to 'assetize' vehicular data on top of their 'ordinary' function as mobility vehicles, enabling automotive firms to envision additional 'rent' from vehicle options and vehicles, often through subscriptions (BM-2025). Ford's BlueCruise driver assistance system, for example, is only available to users through a monthly ($49.99) or annual ($495) subscription service rather than as an integrated feature available at point of purchase (FO-2024b).

Regulatory limits

Regulatory pressures regarding testing incidents, licencing protocols, innovation principles, and civic activism have hampered the development of autonomous vehicles.

Publicly-tested autonomous vehicles have been involved in multiple incidents. These have included fatal crashes (VE-2019), undisclosed involvement in accidents (VE-2023), and robotaxis waking city residents up by 'honking' through the night (VE-2024). Bespoke licencing procedures have been heavily criticised. In San Francisco, dedicated permit routes – the Autonomous Vehicle Passenger Service Programs (CP-2024) – were launched by the California Public Utilities Commission (CPUC), responsible for taxi licences in the city, with a former autonomous vehicle executive casting a decisive vote to allow the expansion of robotaxi tests (WAS-2023). Yet prospective operators have refused to publicly release the permit-required operational data citing confidentiality implications (SF-2023).

Municipal transport authorities like the San Francisco County Transportation Authority (SFCTA) have contended that robotaxi firms need to pursue an 'incremental' approach to innovation, cleared to expand operations only on successful completion of previous milestones. Such an approach, they argued, would 'offer the best path towards public confidence in driving automation and industry success in San Francisco and beyond'. (SF-2023). Citizen groups have exerted pressure on policymakers and engaged in forms of direct action to limit, or stop, autonomous vehicles. This has included union support for banning robotaxis in Boston (BO-2025), and setting fire to Waymo robotaxis during protests in Los Angeles and other cities (GU-2024; NY-2025), with the vehicles perceived as representative of a rising tech-led authoritarianism (Table 1).

AV 2.0: LLMs as overcoming limits

Using Nvidia GTC 2024's Nvidia Drive session as an entry point, the article has so far examined the difficulties faced by an industry always just 'a few years' away from a world of autonomous driving. In these cases, the technical limits of machine vision, for example, have come up against the economic limits of platform business models, whilst the financial limits of assetization have been matched by an unpredictable global regulatory environment.

Table 1. Technical, financial, economic and regulatory limits to autonomous vehicles.

Technical • Need for infinite hard-coded rules, categories, instructions • Object-recognition errors and performance • System interoperability (up/downstream operational 'pipeline') • Maintenance, supervision, and intervention issues (e.g., robotaxis)	**Financial** • LLM/genAI withdrawing capital funds and VC investment • Continuing high capital expenditure • 'Assetization' of vehicles (data extraction, rentierism)
Economic • Market power of established automotive firms • Failed platform model implementation • Rise of Chinese automotive firms (e.g., BYD)	**Regulatory** • Public failures of robotaxis (e.g., Cruise in San Francisco) • Licencing disputes • Collective pushback from interested parties (e.g., municipal transport authorities) • Rising civic activism (e.g., Safe Street Rebel)

In this section, we examine two companies that have claimed to have leveraged the power of general-purpose LLMs to address these challenges. The first is Wayve, with its GAIA 1.0 model, which uses 'video prompting' techniques to generate training data, and the second is Waabi's CoPilot4D, which promises to allow autonomous vehicles to understand their environment better through 'parallel tokenization'.

Video prompting: wayve GAIA 1.0

Wayve describes GAIA 1.0 as a 'new generative AI model' built to create 'realistic driving videos by leveraging video, text and action inputs' (WAY-2023a). By learning to 'predict the subsequent frames in a video sequence' (WAY-2023a) drawn from real-world data collected by Wayve vehicles, GAIA 1.0 'resembles the approach seen in large language models (LLMs)' (WAY-2023a), typically referred to as a Vision Language Model (VLM) (IB-2025). Here the stated task of GAIA 1.0 is to be able to generate potentially infinite numbers of driving scenes and scenarios in order to train a model capable of not only recognising objects in the world, but possible (and likely) interactions – and conflicts – between them (Figure 2).

GAIA 1.0 is presented as more than simply an incremental step in building a machine vision system for use in the domain of autonomous driving. It is a 'game-changing generative AI research model' (WAY-2023a), providing 'new possibilities for innovation in the field of autonomy, enabling enhanced and accelerated training of autonomous driving technology' (WAY-2023b). The 'true marvel' of GAIA 1.0 'lies in its ability to manifest the generative rules that underpin the world we inhabit' (WAY-2023a).

The main LLM innovation here is *video prompting*, making use of foundation models to offer engineers the possibility of generating a huge volume and variety of synthetic video sequences. This claims to be able to offer an entirely new horizon for autonomous vehicle development, with firms like Wayve no longer limited by data they can capture in real-world environments or acquire through other means, like dashcam footage, or even by the expertise of their engineers to design appropriate text-based prompts. The promise

Figure 2. Simulated driving scenes generated by GAIA 1.0 (WAY-2023b).

is further exacerbated by the emphasis on Nvidia's generative AI-ready stack, with Kendall saying that the chip giant's GPUs and software stack are what allows Wayve 'to build billion-parameter models trained on petabytes of data' (NV-2024b).

Accordingly, it seeks to overcome technical, economic, and likely also the regulatory limits of past approaches to autonomous driving. Regarding technical limits, video prompts offer the possibility of generating a near-infinite volume of scenarios, multiplying the diversity of driving scenes and interactions even where 'edge cases' (i.e., rarely occurring incidents) persist. Regarding economic limits, video prompts offer the prospect of reducing costs of data collection down to the price of compute required for model training (Steinhoff, 2024). Whilst certainly not zero, it reduces the ongoing expense of operating a large vehicle fleet for data collection purposes and concomitant regulatory hurdles of permits to run real-world tests in multiple locations.

Yet the use of video prompting to generate synthetic data is not without consequences. As Jacobsen (2021) discusses, synthetic data techniques, such as using GANs, have long been touted by the autonomous vehicle industry as a way of solving a 'hard miles' problem of finding edge case scenarios in real-world driving data, as suggested by the founder of UK autonomous vehicle provider, Oxa, at a previous GTC event in 2021 (Jacobsen, 2021). Despite being supposed to be able to bridge the so-called 'reality gap' between real and simulated worlds, techniques such as 'domain adaptation' (where real images are used to train GANs) or 'domain randomization' (where objects are randomly inserted into simulations) still rely on conventional data (Steinhoff & Hind, 2025), requiring continuous collection.

By comparison, if firms like Wayve decide to use synthetic scenes for video prompting, there is the possibility of 'model collapse', 'a degenerative process whereby models forget the true underlying data distribution' (OX-2024) and begin to perform poorly through further model iterations. Switching entirely to synthetic data generation, in part through using synthetic video prompts, would therefore likely lessen the accuracy of machine vision systems being developed to recognise road users in a wide range of street scenes and driving scenarios. Only through combining real and synthetic data, or 'data accumulation' (OA-2024), might model collapse be avoided entirely (Table 2).

Parallel tokenization: Waabi copilot4D

Waabi describe CoPilot4D as a 'novel world modeling approach' that 'tokenizes sensor observations' to 'unlock the power of GPT-like unsupervised learning for robotic agents' (WAA-2024a). Understood as a 'spatio-temporal transformer for world modeling'

Table 2. Recent innovations to overcome limits to autonomous vehicle development.

Autonomous vehicle firm	Autonomous vehicle innovation	AI technology	LLM innovation	Limits overcome	Possible issues
Wayve	GAIA 1.0	Vision Language Model (VLM)	Video prompting	Technical, economic, regulatory	Continued real world data reliance, model collapse
Waabi	CoPilot4D	World model/ agent forecasting	Parallel tokenization	Technical, financial	Token exponentialism, computational parallelism

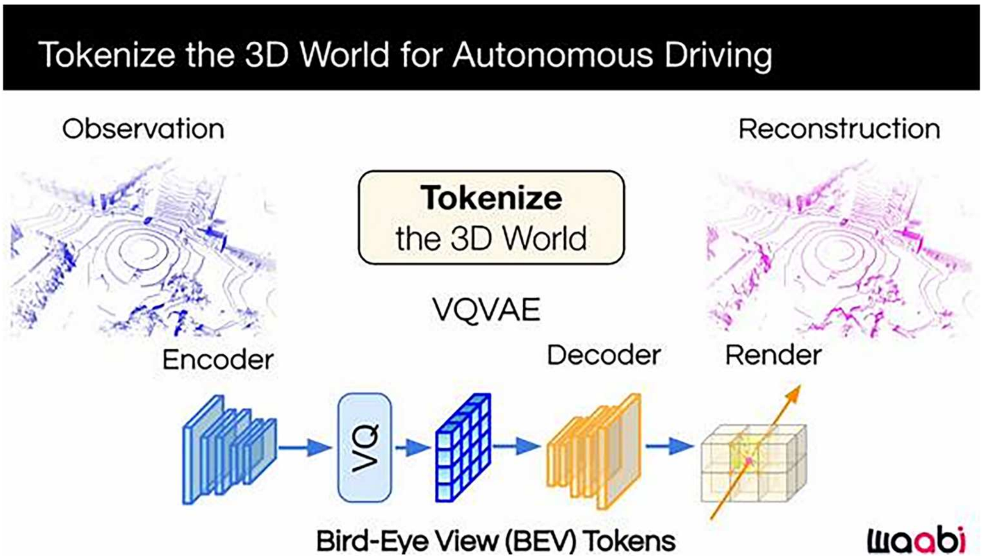

Figure 3. CoPilot4D's 'tokenize everything' approach to the 3D world (WAA-2024a).

(WAA-2024c) – Waabi's CoPilot4D is reported to perform three times better than rival forecasting models (WAA-2024c). Here, the stated task of CoPilot4D is to predict the future state of an autonomous vehicle compared to other agents (road users) in any given scene. Such efforts join the growing attempt to build spatial 'world models' (Mitrokhov, 2024), such as Niantic's Large Geospatial Model (LGM), that condense and extrapolate physical knowledge in a spatial equivalent of LLMs (NI-2024), as exemplified by Niantic's partnership with robot delivery operator Coco Robotics (NI-2026) (Figure 3).

The principal LLM innovation is *parallel tokenization*, designed to apply the concept to objects (i.e., pixels) comprising video footage captured of street scenes and driving scenarios. Similar to the notion of attention in LLMs, the model allows for a much larger context window that carries a persistent 'understanding' of the world and the various entities (other cars, pedestrians) inhabiting it. The intention is that parallel tokenization will be able to offer unsupervised learning in a domain where supervised learning is the norm due to the complexity of the environments being modelled. 'The task', as they consider, 'is essentially about building an unsupervised world model on Lidar sensor observations', alone (WAA-2024c). In doing so, the same model trained in the simulation should be able to extrapolate itself in real-world conditions, including when encountering geographical features not originally found in the training data.

The limits that parallel tokenization is designed to overcome are therefore both technical and financial by generating world models directly from sensor observations. Supervised learning's hard-coded approaches to classifying objects have historically resulted in incorrect object classification and vehicle responses (VE-2019). They also carry substantial labour and data integration costs typically required to build, classify, and maintain training datasets.

The clear issue with applying parallel tokenization to autonomous vehicle object-recognition processes is the sheer volume and variety of object trajectories in any one given scene/scenario. In essence, that the total number of tokens needed to be generated for each frame or sequence of frames is of a much higher degree of magnitude than needed for text corpora – as each scene is likely to involve a certain number of objects of interest (i.e., road users, terrain features, street signs).

A secondary issue of parallel tokenization is therefore the computational capacity needed to perform such processes 'online', in real time, and in parallel. As the designers write again, 'in domains such as autonomous driving, a single observation has tens of thousands of tokens, so parallel decoding of tokens becomes a must.' (WAA-2024c). Parallel approaches have also been commonplace in the development of autonomous vehicles, from parallel sensor systems (MO-2020) to parallel fixed-function and general-purpose processing (MO-2025).

Discussion: driving innovation?

In this discussion, we reflect on the wider implications of our analysis. These can be summarised as follows: the desire of tech firms to reduce human labour, the computational burden of parallel computing, the growing 'assetization' of vehicles, and the social risks of epistemic uncertainty.

Alex Kendall's remarks about the desire for 'non-supervised' learning driven by video prompting are emblematic of aspirations within the industry to reduce reliance on 'invisible' human data work that AI systems require (Gray & Suri, 2019). Yet regulatory issues around the labour conditions of AI 'verifiers' (Tubaro et al., 2020) and the expertise of remote vehicle operators or 'interveners' (Hind, 2022) are only set to intensify in AV 2.0 (Hawkins, 2026). Tech firms are well-versed in these controversies, and have long sought to obscure reliance on human data work, whether for AI verification, remote operation, or social media content moderation (Roberts, 2019). Here, we see a continuation of making workers into 'data assets' (Van Doorn & Badger, 2020), where existing human labour is leveraged into a speculative future in which workers will no longer be needed.

Urtasun's belief in parallel tokenization extends interest in parallel computing. Such methods make further use of Graphics Processing Units (GPUs), first developed by Nvidia for computer game graphics in the 1990s, and sensor-driven smartphone cameras in the 2010s (Mackenzie & Munster, 2019). The application of computational 'parallelism' to world modelling, therefore, will only exacerbate the computational burden of LLMs due to the huge volumes of 3D visual data needing to be processed. Either through the extraction of real-world data or the generation of synthetic data, the approaches will likely exacerbate the environmental impact of AI, dependent upon the construction of more energy-intensive data centres required to enable the growth of AI-dependent technologies (Hogan, 2015; Rone, 2023; Valdivia, 2025).

To scale, both innovations will likely need to contribute to the 'assetization' of vehicles like other digital products (Bernevega & Gekker, 2022; Birch & Muniesa, 2020). Whilst arguably a more recent phenomenon, driven by digital platforms, Forelle and Shapiro (2024) place the shift towards rentier, subscription models in the wider automotive industry within a longer history of its financialization since the 1980s. Like loans and insurance products, 'subscriptions produce positive cash flows that extend in time towards some horizon of futurity' (Forelle & Shapiro, 2024, p. 151), enabling automotive firms to develop entirely new revenue streams in addition to those developed over the last 40 years. Waabi and Wayve point to how general-purpose AI models might extend the trend further into autonomous vehicles: training and inference are 'compute-hungry', which firms might seek to fund through assetization or similar revenue models.

Finally, both companies also considerably underplay the epistemic uncertainty of LLMs. While Urtasun mentions that 'explainability' is important for both technical (efficiency) and regulatory (responsibility) reasons, underlying issues go unacknowledged. The same thing that makes LLMs exceptionally fitting for solving the limits of autonomous driving – their generality – makes them particularly dangerous when applied to real-world driving scenarios. Whereas a 'hallucinating' chatbot might result in a failed student essay, a 'hallucinating' car poses far greater social risk. As Le (2023) notes, LLMs still struggle with 'long-tail' events not present in training data, and are unable to 'maintain coherent performance over longer time horizons' (Backlund & Petersson, 2025, p. 1). Translating these innovations to real-world driving situations will require much more domain-specific research into the social consequences.

Conclusion: driving hype?

As this article has argued, proponents have sought to bring the curtain down on AV 1.0, the consolidated attempt over the last decade to commercialise autonomous vehicles. Yet, rather than disappear entirely, a supposed new dawn has risen – AV 2.0 – driven by LLM-style innovations. Intended as a clean break from well-documented past failures, AV 2.0 has driven a new wave of hope – and hype (Kotliar, 2025). Yet to do so, these stated new technologies still need to tackle a series of limits that have foiled the commercialisation of autonomous vehicles before. As we have argued, these limits – technological, financial, economic, and regulatory – will not be easily overcome, requiring much more than Big Tech bravado to surmount.

The article has examined two innovations of this emerging era, each underpinning the announcement of new technologies by two autonomous vehicle firms, Wayve and Waabi. Despite the stated technical innovations of Wayve's GAIA 1.0 or Waabi CoPilot4D – attempts to develop 'video prompting' and 'parallel tokenization' dependent on LLM techniques – hard limits to autonomous vehicles remain. The innovations discussed in this article had forced autonomous vehicle firms to retreat to the research lab, placing their faith in general-purpose, LLM-style, approaches to drive a new era of prospective commercialisation. At present these bets remain just as provisional as past attempts, requiring continued 'breakthroughs' not only in technical work in novel approaches like video prompting or parallel tokenization but in financial, economic, and regulatory spheres too.

We draw three final conclusions from our analysis. First, it is important to be alert to *discursive claims* of 'general-purposivity' as LLMs and generative AI products work their way into new industries and worlds (Kotliar, 2025). Whilst it is likely that some claims about the general-purposivity of LLMs will continue to be imported 'wholesale' into applied domains, discursive claims also invariably undergo modification, suited to the cultural norms and expectations within each specific industry. Wayve's promise of ditching the costly human aspect of autonomous vehicle training resonates with Big Tech's long-held dreams of automating away a 'pesky' human workforce. As the case of autonomous vehicles demonstrates, 'general' technologies in the lab always require *specification* in the real world.

Accordingly, it is crucial to examine how *techniques* of general-purposivity (Burkhardt & Rieder, 2024) are being developed to suit operational needs and requirements within specific industries and sectors, as part of AI's ongoing 'industrialisation' (van der Vlist et al., 2024). As exemplified by Waabi, the computational demands of new LLM-style methods easily outstrip available resources. In such cases, the innovations behind LLMs do not necessarily translate into new contexts, but require the further reshaping of existing norms, methods, and workflows.

Then lastly, it is useful to consider how the *value* of general-purposivity is likely to be generated in and circulate through particular industries, and the extent to which LLM-style innovations will continue to rely – in one way or another – on the infrastructural power of Big Tech and AI firms (Luitse, 2024). The computational, human, and environmental costs should be taken into account when evaluating – and contesting – inflated claims to the general-purposivity of LLMs, whether in the tech industry or the many new industries being asked, encouraged, or indeed *forced* to adopt them. In other words, it is necessary to figure out who exactly is driving the hype and why.

Notes

1. We recognise that claims made in ArXiv papers – and on similar preprint servers – are not subject to peer review. Even when they do, it is often impossible for other researchers to replicate what they claim due to lack of access to the kinds of resources (be they data or compute) at hand inside the firms. However, as stated in our overall approach, we see this as part of the discursive claims made by the AI companies, similar to their reports and video presentations.
2. A third presenting company, Nuro, was not included in the analysis as they were neither a significant autonomous vehicle firm in AV 1.0 nor have explicitly used the AV 2.0 term. As such they have not contributed to discursive framings of a 'rebooting' of the autonomous vehicle industry across this time period. For similar reasons, two representatives of major automotive companies (Ford and Jaguar) did not present information that warranted inclusion in the case studies.
3. Fellow Chinese automotive firms Geely and SAIC were hit with higher tariffs for allegedly failing to cooperate fully with the European Commission's investigation (RE-2024).

Acknowledgements

The authors would like to thank participants and audience members of the Smart' Technologies panel at the Association of Internet Researchers (AoIR) annual conference in Sheffield (2024) at which this article was first presented. The authors would additionally like to thank Joan Ramon Rodríguez-Amat and Adriana de Rosa Amaral for their invitation to this special issue. A special thanks to Bernhard Rieder who provided helpful feedback on an early version of the conceptual framework, and to the three reviewers who offered insightful and constructive feedback. All errors within remain entirely the authors' own.

Disclosure statement

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

Notes on contributors

Alex Gekker is an Assistant Professor in Digital Research Methods, Media Studies Department, University of Amsterdam. His research incorporates various aspects of digital media, primarily focusing on platforms, infrastructures and interfaces to analyse diverse objects such as data centres, maps, surveillance assemblages, autonomous cars, videogame ecosystems and more. His book Maps as Media (2026) is published by Polity Press.

Sam Hind is Lecturer in Digital Media and Culture at the University of Manchester. His research interests include machine learning cultures, tech innovation and competitions, autonomous driving, and the platformisation of automobility. He is the author of Driving Decisions: How Autonomous Vehicles Make Sense of the World (2024).

ORCID

Alex Gekker http://orcid.org/0000-0001-6042-2086
Sam Hind http://orcid.org/0000-0001-8347-3695

References

Agre, P. (1994). Surveillance and capture: Two models of privacy. *The Information Society*, *10*(2), 101–127. https://doi.org/10.1080/01972243.1994.9960162.

Ameisen, E., Lindsey, J., Pearce, A., Gurnee, W., Turner, N. L., Chen, B., Citro, C., Abrahams, D., Carter, S., Hosmer, B., Marcus, J., Sklar, M., Templeton, A., Bricken, T., McDougall, C., Cunningham, H., Henighan, T., Jermyn, A., Jones, A., … Batson, J. (2025, March 27). Circuit tracing: Revealing computational graphs in language models. *Transformer Circuits Thread*. https://transformer-circuits.pub/2025/attribution-graphs/methods.html

Anthropic. (2026, March 23). Dispatch and computer use. *Anthropic*. https://claude.com/blog/dispatch-and-computer-use

Apple. (2025, July 17). Updates to Apple's on-device and server foundation language models. *Apple* https://machinelearning.apple.com/research/apple-foundation-models-2025-updates

Backlund, A., & Petersson, L. (2025, February 20). Vending-bench: A benchmark for long-term coherence of autonomous agents. *arXiv (Computing Research Repository)*, 1–28. https://doi.org/10.48550/arXiv.2502.15840

Bernevega, A., & Gekker, A. (2022). The industry of landlords: Exploring the assetization of the triple-A game. *Games and Culture*, *17*(1), 47–69. https://doi.org/10.1177/15554120211014151.

Birch, K., & Muniesa, F. (2020). *Assetization: Turning things into assets in technoscientific capitalism*. MIT Press.

Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, *20*(1), 30–44. https://doi.org/10.1080/1369118X.2016.1154086.

Bucher, T. (2018). *If … then: Algorithmic power and politics*. Oxford University Press.

Burkhardt, S., & Rieder, B. (2024). Foundation models are platform models: Prompting and the political economy of AI. *Big Data & Society*, *11*(2), 1–15. https://doi.org/10.1177/20539517241247839.

Centre for Research on Foundation Models (CRFM). (2022). On the opportunities and risks of foundation models. *arXiv (Computing Research Repository)*, 1–214. https://doi.org/10.48550/arXiv.2108.07258

Dalton, C. M. (2015). For fun and profit: The limits and possibilities of google-maps-based geoweb applications. *Environment and Planning A: Economy and Space*, *47*(5), 1029–1046. https://doi.org/10.1177/0308518X15592302.

De Kloet, J. (2026). Longing for an AI winter—beyond techno-utopia, towards a return of the personal. *Global Media and China*, *11*(1), 136–144. https://doi.org/10.1177/20594364261426921.

Douglas Heaven, W. (2024, June 18). Why does AI hallucinate? *MIT Technology Review*. (https://www.technologyreview.com/2024/06/18/1093440/what-causes-ai-hallucinate-chatbots/)https://www.technologyreview.com/2024/06/18/1093440/what-causes-ai-hallucinate-chatbots/

Douglas Heaven, W. (2026, January 12). Mechanistic interpretability. *MIT Technology Review*. (https://www.technologyreview.com/2024/06/18/1093440/what-causes-ai-hallucinate-chatbots/)https://www.technologyreview.com/2026/01/12/1130003/mechanistic-interpretability-ai-research-models-2026-breakthrough-technologies/

Edwards, B. (2024, October 17). Cheap AI 'video scraping' can now extract data from any screen recording. *ArsTechnica*. (https://arstechnica.com/ai/2024/10/cheap-ai-video-scraping-can-now-extract-data-from-any-screen-recording/)https://arstechnica.com/ai/2024/10/cheap-ai-video-scraping-can-now-extract-data-from-any-screen-recording/

Egliston, B., & Carter, M. (2022). Oculus imaginaries: The promises and perils of Facebook's virtual reality. *New Media & Society*, *24*(1), 70–89. https://doi.org/10.1177/1461444820960411

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023, August 21). GPTs are GPTs: An early look at the labor market impact potential of large language models. *arXiv (Computing Research Repository)*, 1–36. https://doi.org/10.48550/arXiv.2303.10130

Ferreira, B. (2026, March 7). Claude Code deletes developers' production setup, including its database and snapshots — 2.5 years of records were nuked in an instant. *Tom's Hardware*. https://www.tomshardware.com/tech-industry/artificial-intelligence/claude-code-deletes-developers-production-setup-including-its-database-and-snapshots-2-5-years-of-records-were-nuked-in-an-instant

Foley, J. (2025, May 29). Nvidia is a $3tn bet on the tokenization of everything. *Financial Times*. (https://www.ft.com/content/0f6b115e-0d2c-4dd7-8aa5-a031b7cd4bf3)https://www.ft.com/content/0f6b115e-0d2c-4dd7-8aa5-a031b7cd4bf3

Forelle, M. C., & Shapiro, A. (2024). Driving into a paywall: The subscriptionization of consumer vehicles. *Engaging Science, Technology, and Society*, *10*(3), 147–180. https://doi.org/10.17351/ests2023.2369

Förster, S. (2023). The bigger the better? *A Peer-Reviewed Journal About*, *12*(1), 120–134. https://doi.org/10.7146/aprja.v12i1.140444

Frankfurt, H. (2005). *On bullshit*. Princeton University Press.

Funk, J. (2019). What's behind technological hype? *Issues in Science and Technology*, *36*(1), 36–42. https://issues.org/behind-technological-hype/

Goldfarb, A., Taska, B., & Teodoridis, F. (2023). Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings. *Research Policy*, *52*(1), 1–25. https://doi.org/10.1016/j.respol.2022.104653

Gray, M. L., & Suri, S. (2019). *Ghost work: How to stop Silicon Valley from building a new global underclass*. Houghton Mifflin Harcourt.

Hawkins, A. J. (2022, October 26). Argo AI, driverless startup backed by Ford and VW, is shutting down. *The Verge*. https://www.theverge.com/2022/10/26/23423998/argo-ai-shut-down-ford-vw-av-selfdriving

Hawkins, A. J. (2026, April 6). Robotaxi companies won't say how often remote operators intervene. *The Verge*. https://www.theverge.com/transportation/907478/robotaxi-remote-assistance-markey-investigation-waymo-tesla

Hern, A. (2024, May 25). Big tech has distracted world from existential risk of AI, says top scientist. (https://www.theguardian.com/technology/article/2024/may/25/big-tech-existential-risk-ai-scientist-max-tegmark-regulations)https://www.theguardian.com/technology/article/2024/may/25/big-tech-existential-risk-ai-scientist-max-tegmark-regulations

Hicks, M. T., Humphries, J., & Slater, J. (2024). ChatGPT is bullshit. *Ethics and Information Technology*, *26*(2), 1–10. https://doi.org/10.1007/s10676-024-09775-5

Hind, S. (2022). Making decisions: The normal interventions of nissan 'mobility managers'. *Mobilities*, *17*(4), 467–483. https://doi.org/10.1080/17450101.2021.1988682

Hind, S. (2024a). *Driving decisions: How autonomous vehicles make sense of the world*. Palgrave Macmillan.

Hind, S. (2024b). From united steel to waymo: Industrializing simulation. *AI & Society*, *40*(4), 2015–2028. https://doi.org/10.1007/s00146-024-02051-6

Hogan, M. (2015). Data flows and water woes: The Utah data center. *Big Data & Society*, *2*(2), 1–12. https://doi.org/10.1177/2053951715592429

Jacobsen, B. (2021). Machine learning and the politics of synthetic data. *Big Data & Society*, *10*(1), 1–12. https://doi.org/10.1177/20539517221145372

Jasanoff, S., & Kim, S.-H. (2013). Sociotechnical imaginaries and national energy policies. *Science as Culture*, *22*(2), 189–196. https://doi.org/10.1080/09505431.2013.786990

Ji, J., Qiu, T., Chen, B., Zhang, B., Lou, H., Wang, K., Duan, Y., He, Z., Vierling, L., Hong, D., Zhou, J., Zhang, Z., Zeng, F., Dai, J., Pan, X., Yee Ng, K., O'Gara, A., Xu, H., Tse, B., … Gao, W. (2025, April 4). AI alignment: A comprehensive survey. *arXiv (Computing Research Repository)*, 1–105. https://doi.org/10.48550/arXiv.2310.19852

Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., & Amodei, D. (2020, January 23). Scaling laws for neural language models. *arXiv (Computing Research Repository)*, 1–30. https://doi.org/10.48550/arXiv.2001.08361

Korosec, K., & Bellan, R. (2023, November 19). Cruise co-founder and CEO Kyle Vogt resigns. TechCrunch. https://techcrunch.com/2023/11/19/cruise-co-founder-and-ceo-kyle-vogt-resigns

Kotliar, D. (2025). Can't stop the hype: Scrutinizing AI's realities. *Information, Communication & Society*, *29*(3), 828–849. https://doi.org/10.1080/1369118X.2025.2531165

Le, J. (2023, March 31). Foundation models are going multimodal. *Twelve Labs*. (https://www.twelvelabs.io/blog/foundation-models-are-going-multimodal)https://www.twelvelabs.io/blog/foundation-models-are-going-multimodal

Li, C., Gan, Z., Yang, Z., Yang, J., Li, L., Wang, L., & Gao, J. (2023, September 18). Multimodal foundation models: From specialists to general-purpose assistants. *arXiv (Computing Research Repository)*, 1–119. https://doi.org/10.48550/arXiv.2309.10020

Luitse, D. (2024). Platform power in AI: The evolution of cloud infrastructures in the political economy of artificial intelligence. *Internet Policy Review*, *13*(2), 1–44. https://doi.org/10.14763/2024.2.1768.

Luitse, D., & Denkena, W. (2021). The great transformer: Examining the role of large language models in the political economy of AI. *Big Data & Society*, *8*(2), 1–14. https://doi.org/10.1177/20539517211047734.

Mackenzie, A., & Munster, A. (2019). Platform seeing: Image ensembles and their invisualities. *Theory, Culture & Society*, *36*(5), 3–22. https://doi.org/10.1177/0263276419847508.

Metz, C., & Griffith, E. (2020, May 12). This was supposed to be the year driverless cars went mainstream. *The New York Times*. https://www.nytimes.com/2020/05/12/technology/self-driving-cars-coronavirus.html

Mitrokhov, K. (2024). Between world models and model worlds: On generality, agency, and worlding in machine learning. *AI & Society*, *40*(6), 5087–5099. https://doi.org/10.1007/s00146-024-02086-9.

Muldoon, J., Cant, C., Wu, B., & Graham, M. (2024). A typology of artificial intelligence data work. *Big Data & Society*, *11*(1), 1–13. https://doi.org/10.1177/20539517241232632.

Narayan, D. (2022). Platform capitalism and cloud infrastructure: Theorizing a hyper-scalable computing regime. *Environment and Planning A: Economy and Space*, *54*(5), 911–929. https://doi.org/10.1177/0308518X221094028.

Narayanan, A., & Kapoor, S. (2025). Why an overreliance on AI-driven modelling is bad for science. *Nature*, *640*(8058), 312–314. https://doi.org/10.1038/d41586-025-01067-2.

OpenAI (2025a, March 4). GPT-4 technical report. *arXiv (Computing Research Repository)*, 1–100. https://doi.org/10.48550/arXiv.2303.08774

OpenAI. (2025b, no date). Platform: Models. *OpenAI*. https://platform.openai.com/docs/models

OpenAI. (2025c, January 23). Introducing Operator. *OpenAI*. https://openai.com/index/introducing-operator//

Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media*. Yale University Press.

Rone, J. (2023). The shape of the cloud: Contesting date centre construction in north holland. *New Media & Society*, *26*(10), 5999–6018. https://doi.org/10.1177/14614448221145928

Roth, E. (2024, February 27). Apple's electric car project is dead. *The Verge*. https://www.theverge.com/2024/2/27/24084907/

Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2025). AI agents vs. Agentic AI: A conceptual taxonomy, applications and challenges. *arXiv (Computing Research Repository)*, 1–30. https://doi.org/10.48550/arXiv.2505.10468

Schüll, N. D. (2014). *Addiction by design: Machine gambling in Las vegas*. Princeton University Press.

Shalev-Shwartz, S., Shammah, S., & Shashua, A. (2018, October 27). On a formal model of safe and scalable self-driving cars. *arXiv (Computing Research Repository)*, 1–37. https://doi.org/10.48550/arXiv.1708.06374

Stackpole, B. (2024, August 6). The impact of generative AI as a general-purpose technology. *MIT Sloan*. (https://mitsloan.mit.edu/ideas-made-to-matter/impact-generative-ai-a-general-purpose-technology)https://mitsloan.mit.edu/ideas-made-to-matter/impact-generative-ai-a-general-purpose-technology

Steinhoff, J. (2024). Toward a political economy of synthetic data: A data-intensive capitalism that is not a surveillance capitalism? *New Media & Society*, *26*(6), 3290–3306. https://doi.org/10.1177/14614448221099217.

Steinhoff, J., & Hind, S. (2025). Simulation and the reality gap: Moments in a prehistory of synthetic data. *Big Data & Society*, *12*(1), 1–14. https://doi.org/10.1177/20539517241309884.

Steyerl, H. (2023). Mean images. *New Left Review*, *140/141*, 82–97. https://newleftreview.org/issues/ii140/articles/hito-steyerl-mean-images

Trajtenberg, M. (2018). *AI as the next GPT: A political-economy perspective*. National Bureau of Economic Research. https://doi.org/10.3386/w24245

Tubaro, P., Casilli, A. A., & Coville, M. (2020). The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence. *Big Data & Society*, *7*(1), 205395172091977–12. https://doi.org/10.1177/2053951720919776.

Valdivia, A. (2025). The*supply chain capitalism of AI*: A call to (re)think algorithmic harms and resistance through environmental lens. *Information, Communication & Society*, *28*(12), 2118–2134. https://doi.org/10.1080/1369118X.2024.2420021.

van der Vlist, F., Helmond, A., & Ferrari, F. (2024). Big AI: Cloud infrastructure dependence and the industrialisation of artificial intelligence. *Big Data & Society*, *11*(1), 1–16. https://doi.org/10.1177/20539517241232630.

Van Doorn, N., & Badger, A. (2020). Platform capitalism's hidden abode: Producing data assets in the gig economy. *Antipode*, *52*(5), 1475–1495. https://doi.org/10.1111/anti.12641.

Waymo. (2025, December 9). Demonstrably safe AI for autonomous driving. Waypoint: The Official Waymo Blog. https://waymo.com/blog/2025/12/demonstrably-safe-ai-for-autonomous-driving/

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., … Rush, A. (2020). Transformers: State-of-the-Art natural language processing. In Q. Liu, & D. Schlangen (Eds.), *Proceedings of the 2020 conference on empirical methods in natural language processing: System demonstrations* (pp. 38–45). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-demos.6.

Wong, M. (2024, October 17). The AI boom has an expiration date. *The Atlantic*. (https://www.theatlantic.com/technology/archive/2024/10/agi-predictions/680280/)https://www.theatlantic.com/technology/archive/2024/10/agi-predictions/680280/

Woolgar, S. (1999). A new theory of innovation? *Prometheus*, *16*(4), 441–452. https://doi.org/10.1080/08109029808629293.

Appendix

No.	Citation	Source	Type	Link
1	BM-2025	BMW	Business	(2025, no date). Driving Assistant Plus. https://www.bmw.co.uk/en/shop/ls/dp/DrivingAssistant_Plus_SFA_gb
2	BO-2025	Boston City Council	Political	(2025, October 14) City Council Committee on Government Operations Hearing on Docket #1432 https://www.boston.gov/public-notices/16486776
3	BY-2022	BYD	Business	(2022, December 29). Hello! We are BYD! https://www.byd.com/uk/blog/Hello-we-are-BYD
4	CP-2024	California Public Utilities Commission (CPUC)	Political	(2024, no date). Autonomous Vehicle Passenger Service Programs. https://www.cpuc.ca.gov/regulatory-services/licensing/transportation-licensing-and-analysis-branch/autonomous-vehicle-programs
5	CN-2024	CNN	Media	(2024, September 13). Biden finalises increases to some of Trump's China tariffs. https://edition.cnn.com/2024/09/13/politics/china-tariffs-biden-trump/index.html
6	EU-2024	European Commission	Political	(2024, October 29). EU imposes duties on unfairly subsidised electric vehicles from China while discussions on price undertakings continue. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_5589
7	FI-2024a	Financial Times	Media	(2024, December 23). Honda and Nissan unveil plan for $58bn merger by 2026. https://www.ft.com/content/dbad4d4d-effc-4926-b2c7-c573d5d6deca
8	FI-2024b			(2024, January 4). Wang Chuanfu, the driving force behind BYD's rise. https://www.ft.com/content/22527628-733e-4188-9678-ab3210fdb1ba
9	FO-2024a	Ford	Business	(2024a, no date). Ford BlueCruise: Bringing you hands-free motorway driving. https://www.ford.co.uk/technology/driving-assistance/ford-bluecruise
10	FO-2024b			(2024b, October 1). Ford lowers BlueCruise annual/monthly plan pricing, simplifies access to hands-free highway driving. https://media.ford.com/content/fordmedia/fna/us/en/news/2024/10/01/ford-lowers-bluecruise-annual–monthly-plan-pricing.html
11	GU-2023	The Guardian	Media	(2023, April 11). TechScape: The end of the 'free money' era'. https://www.theguardian.com/technology/2023/apr/11/techscape-zirp-tech-boom
12	GU-2024			(2024, February 12). Driverless taxi vandalised and set on fire in San Francisco's Chinatown. https://www.theguardian.com/us-news/2024/feb/12/waymo-car-fire-san-francisco
13	IB-2025	IBM	Business	(2025, February 25). What are vision language models (VLMs)? https://www.ibm.com/think/topics/vision-language-models
14	IE-2025	International Energy Agency (IEA)	Political	(2025, May 14). Global EV outlook 2025. https://www.iea.org/reports/global-ev-outlook-2025/trends-in-the-electric-car-industry-3
15	IN-2025	The Independent	Media	(2025, March 18). Nvidia CEO Jensen Huang discusses AI's future at GTC 2025. https://www.independent.co.uk/news/nvidia-general-motors-san-jose-woodstock-b2717492.html
16	MO-2020	Mobileye	Business	(2020, December 9). Autonomous-vehicle safety demands true redundancy. https://www.mobileye.com/blog/av-safety-demands-true-redundancy/
17	MO-2025			(2025, February 27). Mobileye's new ECU series: Modularity from ADAS to AV. https://www.mobileye.com/blog/modularity-from-adas-to-av/
18	NI-2024	Niantic	Business	(2024, November 12). Building a large geospatial model to achieve spatial intelligence. https://nianticlabs.com/news/largegeospatialmodel/
19	NI-2026			(2026, October 10). Niantic Spatial Partners with Coco Robotics to Accelerate the Future of Autonomous

				Delivery. https://www.nianticspatial.com/en/blog/coco-robotics
20	NV-2024a	Nvidia	Business	(2024, March 18). Driving innovation: Generative AI and industrial digitisation in the automotive industry. https://www.nvidia.com/en-us/on-demand/session/gtc24-s62380/
21	NV-2024b			(2024, May 29). Riding the Wayve of AV 2.0, driven by generative AI. https://blogs.nvidia.com/blog/wayve-generative-ai/
22	NY-2025	The New York Times	Media	(2025, June 9). Waymo halts service in downtown Los Angeles after its taxis are set on fire. https://www.nytimes.com/2025/06/09/us/waymo-la-protests-vehicles-fire.html
23	OA-2024	OpenAI	Business	(2024, April 29). Is model collapse inevitable? Breaking the curse of recursion by accumulating real and synthetic data. https://doi.org/10.48550/arXiv.2404.01413
24	OX-2024	University of Oxford	University	(2024, July 24). AI models collapse when trained on recursively generated data. https://doi.org/10.1038/s41586-024-07566-y
25	RE-2024	Reuters	Media	(2024, May 3). EU says information from three Chinese EV makers insufficient. https://www.reuters.com/business/autos-transportation/eu-says-information-three-chinese-ev-makers-insufficient-2024-05-03/
26	SF-2023	San Francisco Country Transportation Authority (SFCTA)	Political	(2023, no date). Re: Protest of Cruise LLC Tier 2 advice letter. https://www.cpuc.ca.gov/-/media/cpuc-website/files/uploadedfiles/cpucwebsite/content/licensing/autovehicle/cruise-al-02-protests-and-responses-revised.zip
27	ST-2025	Statista	Statistics	(2025, February 10). Global automobile market share in 2024, by brand. https://www.statista.com/statistics/316786/global-market-share-of-the-leading-automakers/
28	UB-2023	Uber	Business	(2023, October 26). Autonomous rides are arriving on Uber with Waymo. https://www.uber.com/us/en/newsroom/waymo-on-uber/
29	UB-2026			(2026, March 26). Verne, Pony.ai and Uber partner to launch Europe's first commercial robotaxi service. https://investor.uber.com/news-events/news/press-release-details/2026/Verne-Pony-ai-and-Uber-Partner-to-Launch-Europes-First-Commercial-Robotaxi-Service/
30	VE-2019	The Verge	Media	(2019, November 20). The world's first robot car death was the result of human error – and it can happen again. https://www.theverge.com/2019/11/20/20973971/uber-self-driving-car-crash-investigation-human-error-results
31	VE-2023			(2023, October 24). California DMV suspends Cruise's robotaxi permit 'effective immediately'. https://www.theverge.com/2023/10/24/23930629/california-dmv-suspends/cruise-robotaxi-permit-safety
32	VE-2024			(2024, August 11). A nightly Waymo robotaxi parking lot honkfest is waking San Francisco neighbours. https://www.theverge.com/2024/8/11/24218134/waymo-parking-lot-livestream-honking-4am-san-francisco
33	WAA-2024a	Waabi	Business	(2024, no date). Copilot4D: Learning unsupervised world models for autonomous driving via discrete diffusion. https://waabi.ai/copilot-4d/
34	WAA-2024b			(2024, March 19). Accelerating the new era of autonomous vehicles with generative AI. https://www.nvidia.com/en-us/on-demand/session/gtc24-s62621/
35	WAA-2024c			(2024, April 1). Copilot4D: Learning unsupervised world models for autonomous driving via discrete diffusion. https://doi.org/10.48550/arXiv.2311.01017
36	WAS-2023	Washington Post	Media	(2023, August 8). California just opened the floodgates for self-driving cars. https://www.washingtonpost.com/technology/2023/08/10/san-francisco-robotaxi-approved-waymo-cruise/

37	WAY-2023a	Wayve	Business	(2023, June 17). Introducing GAIA-1: A cutting-edge generative AI model for autonomy. https://wayve.ai/thinking/introducing-gaia1/
38	WAY-2023b			(2023, September 29). GAIA-1: A generative world model for autonomous driving. https://doi.org/10.48550/arXiv.2309.17080
39	WAY-2024			(2024, March 19). How LLMs and Generative AI will enhance the way we experience self-driving cars. https://www.nvidia.com/en-us/on-demand/session/gtc24-s62464/
40	WE-2024	World Economic Forum (WEF)	Political	(2024, September 18). Surviving the tech market correction: What can global venture capitalists learn from Europe? https://www.weforum.org/stories/2024/09/tech-reset-what-can-global-venture-capitalists-learn-from-europe/
41	WI-2026	Wired	Media	(2026, March 26). The Sony Afeela was doomed to fail. https://www.wired.com/story/sony-afeela-was-doomed-to-fail/