This paper was presented at the XXXVII ISPIM Innovation Conference, held in Granada, Spain, on 07 June to 10 June 2026. ISBN 978-952-65771-2-8.

Identifying emerging autonomous driving trends using deep learning topic modelling on real-world fair data

Sarah Maier*

Technische Universität Berlin, Hardenbergstraße 40A, 10623 Berlin, Germany.
Email: s.maier@campus.tu-berlin.de
*Corresponding author

Jan Kratzer

Technische Universität Berlin, Hardenbergstraße 40A, 10623 Berlin, Germany.
Email: jan.kratzer@tu-berlin.de

Abstract: The widespread adoption of autonomous vehicles (AVs) is expected to transform societal dynamics within the foreseeable future. Therefore, it is essential to identify emerging trends in AV technology research and development, as this enables industry practitioners and policymakers to forecast and proactively adjust their strategies, future development investments and regulatory frameworks. Since traditional topic modeling techniques are either time-consuming or limited in data access and in their ability to capture semantic relationships, this study introduces a novel, deep learning-based approach to topic modeling. Using real-world technology fair data guarantees a comprehensive data pool with state-of-the-art technology. The findings revealed seven macro-topics (MTs) and a four-layered modular architecture that is consistent with digital innovation, reflecting the industry's shift towards software-driven, service- and platform-based structures, rather than hardware-driven ones. Furthermore, the successful deployment not only validates the proposed methodology, but also yields a streamlined, transferable workflow for future trend analysis across domains.

Keywords: BERTopic, topic modeling, deep learning, technology forecasting, autonomous vehicle, digital innovation, trend analysis, layered modular architecture

1 Introduction

Advancements in automotive technology have led to improvements in travel comfort, safety, space, and general speed and distance. Current standard products are equipped with a number of sensors that can regulate speed, apply the brakes in an emergency, enable parking and monitor the area around the vehicle, as well as offering many more features. Integrating autonomous mobility solutions into the automotive sector is the next step in research and development. These solutions are already being tested and are partly in use on the roads. Recent discourse emphasises the impact that autonomous vehicles

(AVs) are expected to have on urban mobility, while also highlighting the challenges associated with this disruptive technology (Cohen *et al.*, 2020; Mattas, Botzoris and Papadopoulos, 2021; Bissell *et al.*, 2020). Possible benefits include reduced carbon emissions and travel time, as well as giving drivers the opportunity to engage in other activities instead of spending long periods of time on manual driving. AVs are also expected to significantly enhance the overall travel experience (Stead and Vaddadi, 2019). Despite these obvious advantages, a number of unresolved challenges remain. These include legal, ethical and safety concerns, particularly given the ongoing evolution of legislation and the potential human rights implications of accidents involving AV (Alqahtani, 2025). Another challenge lies in technological developments, as AVs depend on a complex, multi-layered system integrating advanced sensor technology, communication networks and vehicular computing. Some of these technologies go beyond state-of-the-art developments and network capabilities (Yao *et al.*, 2020). At the current stage, AVs operate through an intricate network of sensors, including radar, LiDAR, cameras and ultrasonic sensors. Based on these sensors, machine learning, particularly in the fields of computer vision and deep learning, enables these systems to recognise objects, assess road conditions and predict the movements of other entities within their operational field (Fontan-Costas *et al.*, 2025). Furthermore, V2X (vehicle-to-everything) communication systems enable the data exchange between vehicles and with the surrounding environment. However, the effectiveness of these communication systems is limited by the general network quality and the technological capabilities of all traffic participants (Hakak *et al.*, 2022; Yang *et al.*, 2026). Ultimately, the high cost of sensors, unresolved safety issues resulting from the limitations of current AI models and the lack of uniform regulatory standards across jurisdictions remain critical obstacles to large-scale deployment (Alqahtani, 2025). Nevertheless, as Figure 1 shows, the estimated increase in the availability of AVs is tremendous.

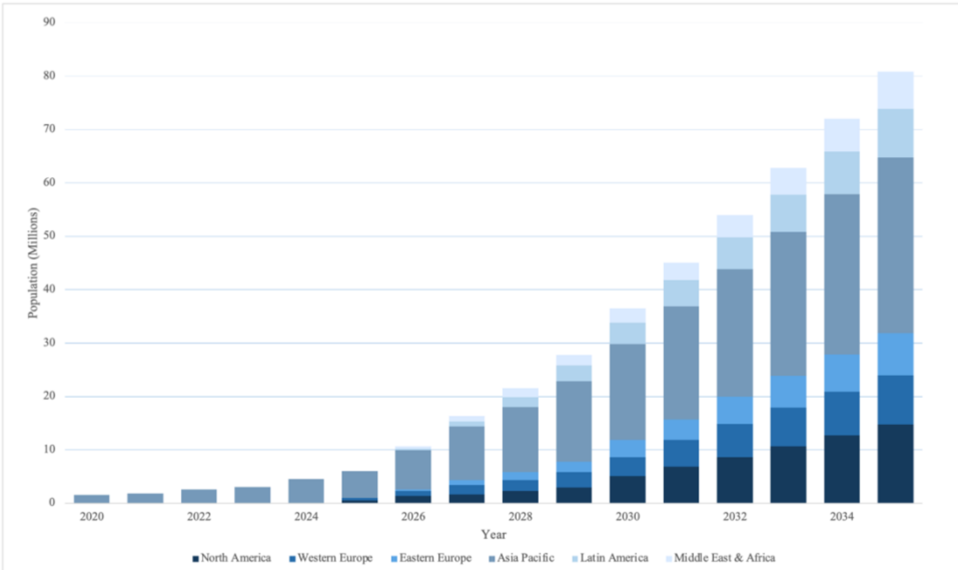

Figure 1 Estimated and expected availability of AVs in several regions worldwide (adapted from Alqahtani, 2025)

In this study, we examine current trends in AV technology research and development using a new, unique method. Since technology fairs often provide an early opportunity to showcase new and near-market technologies, we gathered material for this study at a leading automotive fair in Munich. We visited 93 exhibitor booths at the IAA Summit and the IAA Open Space 2025. This material was used to run a deep learning-based topic modeling approach to identify the latest trends in AV technology development. Our approach opens up new methodologies for technology forecasting and trend analysis, which until now has mainly relied on bibliometric and patent analysis (Daim *et al.*, 2006), systematic literature analysis, surveys, and various data analytics techniques, including AI-driven analytical models (Tsai *et al.*, 2023; Reis *et al.*, 2025). However, none of the technology forecasting attempts so far have focused on technology fairs, which are at the heart of new technology presentation. Investigating these events and collecting the latest information ensures a comprehensive database and enables state-of-the-art technology forecasting and trend identification. This study outlines our approach and presents results on trends in the development of AV technology, making a twofold contribution to current literature on the subject. First, it presents the latest technological trends in AV research and development. This information is tremendously important for automotive producers and suppliers, as it enables them to forecast and plan for the future development and widespread adoption of automotive driving technology. In addition, policymakers can profit from these insights by adjusting strategic policies and issues around regulatory frameworks accordingly. Second, we introduce an innovative methodology that enables practitioners and scientists to swiftly and effectively identify emerging technology trends and predict technological advancements using AI-driven analytics and a unique visual dataset.

2 Methodology

This study identifies emerging trends through a deep learning-based topic modeling approach, contributing to the evolution of technology forecasting methodologies. The methodology section is divided into three sequential steps: database construction; introduction of topic modeling; and application through data coding.

Database

This study uses the deep learning-based topic modeling approach, BERTopic, to analyse data gathered at IAA Mobility (Internationale Automobil-Ausstellung) 2025 in Munich, Germany (Verband der Automobilindustrie e. V., 2025). IAA Mobility is widely considered to be one of the world’s largest mobility platforms. The database consists of visual sources collected from 93 exhibitor booths during 9–10 September 2025 at both the IAA Summit and the IAA Open Space. The list of exhibitors was compiled using the “autonomous driving” filter on the official IAA website (Verband der Automobilindustrie e. V., 2025). Out of all the exhibitors, 132 identified autonomous driving as a key topic. These 132 exhibitors were then subjected to an expert evaluation, resulting in the exclusion of 39 exhibitors based on the following criteria: provider of automotive parts (not specific to autonomous vehicles); electric vehicle charging service; testing service (not specific to autonomous vehicles); and provider specialised in production automation. In addition, consulting services were excluded based on revenue and affinity with

This paper was presented at the XXXVI ISPIM Innovation Conference, held in Bergen, Norway on 15 June to 18 June 2025. ISBN 978-952-65069-9-9.

autonomous vehicles. Following this evaluation, 93 exhibitors remained in the database. A total of 380 visual sources were collected by visiting the 93 exhibitor booths at the IAA Summit and the IAA Open Space. These included 240 photographs taken on site and 140 flyers obtained from exhibitors as physical copies or via a provided link.

Topic modeling

The deep learning-based topic modeling approach, BERTopic, was used to analyse the visual sources (Grootendorst, 2022). From a technical perspective, "topic modeling algorithms are statistical methods that analyse the words of the original texts to discover the themes that run through them, how those themes are connected to each other, and how they change over time" (Blei, 2012). Traditional Latent Dirichlet Allocation (LDA) topic modeling can struggle to capture contextual nuances and semantic relationships between words. To overcome these limitations, this study uses BERTopic, a deep learning-based topic modeling method that uses contextual embeddings to derive coherent topic representations (Grootendorst, 2022). BERT stands for Bidirectional Encoder Representations from Transformers (Devlin *et al.*, 2018). Previous applications of the BERTopic model across different domains have demonstrated its efficiency and precision in identifying hidden themes and tracking emerging trends (Ebadi, Auger and Gauthier, 2022). The process involves the following steps:

The Sentence-BERT framework converts sentences and paragraphs into dense vector representations using pre-trained language models (Reimers and Gurevych, 2019). The semantic properties of these embeddings allow texts with similar meanings to be positioned close to each other in the vector space (Grootendorst, 2022). To facilitate efficient clustering, the dimensionality of document embeddings is then reduced using UMAP (McInnes *et al.*, 2018; Allaoui, Kherfi and Cheriet, 2020). This approach makes clustering computationally feasible while preserving the manifold structure of the document space. The resulting embeddings are then clustered using HDBSCAN, which identifies noise as outliers and enhances topic quality (McInnes, Healy and Astels, 2017; Campello, Moulavi and Sander, 2013). In the final step, each cluster is assigned a topic representation based on its cluster-word distribution (Grootendorst, 2022). A class-based variation of TF-IDF is therefore applied, assessing word importance across clusters rather than individual documents. This results in coherent topic representations.

Data coding

Prior to the implementation of BERTopic, the database was subjected to a series of preprocessing steps. First, text was extracted from the visual sources using optical character recognition and saved as Markdown files. Formatting and structural noise were removed (e.g. HTML comments, Markdown syntax, URLs, special characters), resulting in clean, plain text. Second, quality control was carried out. Documents containing fewer than 30 characters were excluded due to their limited analytical value. Conversely, documents exceeding 15,000 characters were truncated to retain 60% of the beginning and 40% of the end in order to preserve key contextual information. Moreover, a minimum threshold of 35% readable alphabetic tokens was applied to filter out low-quality content. Ultimately, vocabulary control was introduced through an Excel-based whitelist and blacklist file. The whitelist protects domain-specific terms, and the blacklist removes non-informative terms to enhance topic clarity.

After preprocessing, BERTopic was applied to perform topic modeling. The model was set to a maximum of 20 topic clusters (nr_topics=20) in order to strike a balance between thematic granularity and interpretability, bearing in mind the size of the database. The process began with document embedding. The semantic meaning of each document was captured by converting it into a dense 384-dimensional vector using the paraphrase-multilingual-MiniLM-L12-v2 Sentence-Transformer model. This multilingual model was selected in order to accommodate the mixed-language nature of the database. Thus, documents with similar content were assigned to similar vector representations, regardless of their source language. UMAP then reduced the embeddings from 384 dimensions to five while preserving the neighbourhood structure between documents (n_neighbours=12). Subsequently, HDBSCAN clustered the reduced embeddings, setting a minimum size of six documents. The resulting clusters represent potential topics (topic labels have not yet been assigned at this stage). As part of the vectorisation process, the documents were tokenised using CountVectorizer, which created a vocabulary through frequency-based filtering. This included terms that occurred in at least three documents (min_df = 3) and in no more than 85% of documents (max_df = 0.85), thereby removing both rare outliers and common, non-informative terms. To further reduce non-informative terms, blacklisted terms were excluded. Unigram and bigram words were captured to preserve technical expressions (ngram_range=[1,2]).

Each cluster was assigned a topic label by undergoing processes at two levels. First, c-TF-IDF extracted the most distinctive keywords for each topic cluster based on the vectorised vocabulary. At the second level, the Claude API was used to generate descriptive topic labels from three contextual inputs: the c-TF-IDF extracted keywords, representative document samples and domain-specific terms by cross-referencing the keywords with the whitelist. A structural prompt instructed the LLM (large language model) to generate descriptive topic labels with three to seven words for each topic cluster. Moreover, short cluster descriptions were created based on the topic labels and adaptive document sampling (three to five documents, depending on the size of the cluster). Finally, macro-topics (MTs) were created, taking into account the intertopic distance and their thematic coherence. LLM descriptions were added to each macro-topic based on the topic keywords, labels and descriptions to provide clear insights.

3 Findings

This section presents the key results and current trends in the research and development of AV technology, as derived from the 93 exhibitor booths at IAA Mobility 2025. It is divided into two complementary segments: the descriptive analysis examines the structural composition of the dataset, while the autonomous driving macro-topics reveal recurring thematic patterns.

Descriptive analysis

The findings include a descriptive analysis of the 93 selected exhibiting companies, ordered by size and drawn from the business intelligence databases Dealroom and Orbis.

This paper was presented at the XXXVI ISPIM Innovation Conference, held in Bergen, Norway on 15 June to 18 June 2025. ISBN 978-952-65069-9-9.

This reveals their geographic footprint, business profile, scale and maturity, as well as their ownership and funding structure.

The geographic analysis at the regional level revealed that around half of the exhibitors are headquartered in Europe (46%), with the remaining exhibitors split between Asia (34%) and North America (20%). At a national level, Germany, the United States and China dominate, together representing 61% of all exhibitors (Figure 2). Germany takes the lead position, with a higher distribution in the south of the country.

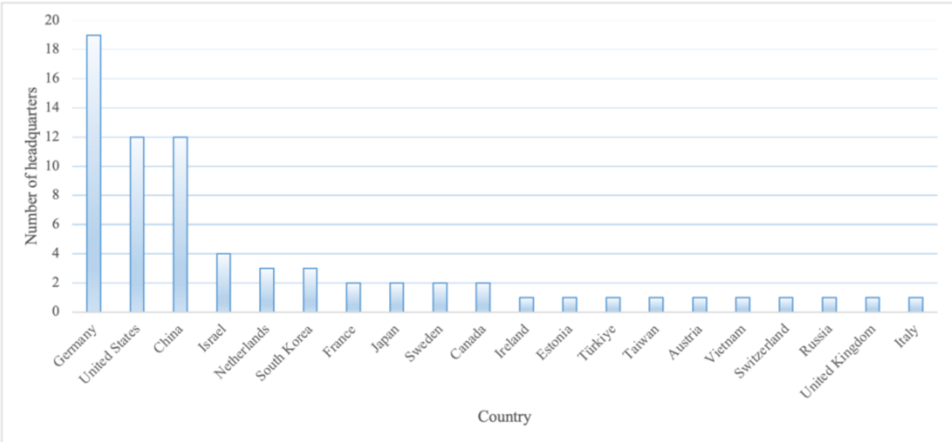

Figure 2 Global company presence by country

The business profile analysis covers the NACE (Nomenclature générale des Activités économiques dans les Communautés Européennes) classification, as well as technology, client focus, revenue model and income stream. NACE is a statistical classification of economic activities developed in the European Union to enhance comparability worldwide (Eurostat, 2025). Each exhibitor has one designated NACE classification. While manufacturing remains dominant, the NACE distribution reveals the technology-driven nature of the exhibitors, with computer programming accounting for 34% (Table 1). When the technologies were examined in more detail, the top three were identified as deep tech (20%), artificial intelligence (16%) and autonomous and sensor tech (13%) (Figure 3). Figure 3 shows how technology associations are distributed among exhibitors, measured by frequency of occurrence. An exhibitor can be associated with multiple technologies.

Table 1 NACE classification of exhibitors (ordered by frequency of occurrence)

NACE classification	*% distribution*
Computer programming activities	34
Manufacture of electronic components	17
Manufacture of motor vehicles	11
Manufacture of communication equipment	11
Business and other management consultancy activities	11
Manufacture of other parts and accessories for motor vehicles	9
Research and experimental development on natural sciences and engineering	5
Legal activities	2

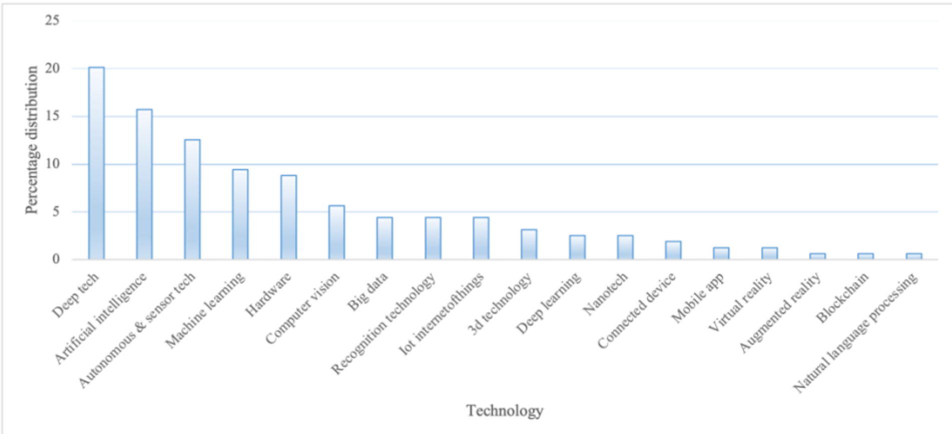

Figure 3 Technology types across exhibitors (ordered by frequency of occurrence)

The majority of exhibitors operate in the B2B (business-to-business) segment (79%), followed by those in the dual B2B/B2C segment (14%), and those in the B2C (business-to-customer) segment (7%). According to the revenue model, around 60% of the exhibitors are in the manufacturing sector, while the remaining 40% are in the software-as-a-service sector. Subscription (46%) and commission-based models (36%) dominate the distribution of income streams, with selling one's own inventory being comparatively low (18%). Taken together, these figures suggest that the industry is transitioning towards service- and platform-based solutions, with even manufacturing companies generating income through services.

When measuring company scale and maturity based on employee numbers, a wide range is revealed among the exhibitors. While the average company employs 33,644 people, the median company size is just 359. Thus, the autonomous driving ecosystem largely comprises medium-sized and emerging companies, as well as a few dominant players. The age of the companies confirms the right-skewed distribution, with a higher average age (34) and a lower median age (22).

The grant structure indicates that the industry is in a state of transition, displaying an innovative, growth-oriented and capital-intensive market (Table 2). The substantial investor interest in companies operating within the autonomous driving ecosystem, as indicated by the high share of venture capital (32%), accelerator (16%) and angel (11%) grants, is driven by their perceived technological advantage (Metrick and Yasuda, 2021, p. 13). Furthermore, the presence of risk capital indicates investor confidence in the ability of new companies to outperform established competitors and scale up rapidly within a large addressable market (Metrick and Yasuda, 2021, pp. 6, 13; Cumming and Johan, 2013, p. 3). The significant proportion of public companies (24%) suggests that the autonomous driving ecosystem is characterised not only by start-ups, but also by the active involvement of established companies. In addition, the low proportion of bootstrapped grants (4%) highlights that this is a capital-intensive market.

Table 2 Types of grants awarded to exhibitors (ordered by frequency of occurrence)

Grant type	*% distribution*
Venture capital	32
Public	24
Accelerator	16
Angel	11
Private equity	7
Subsidiary	6
Bootstrapped	4
Crowdfunded	1

Autonomous driving macro-topics

When the BERTopic model was employed to analyse the 380 visual sources from IAA Mobility 2025, 307 of them passed the quality filter (with a minimum of 35% readable alphabetic tokens and a minimum 30 characters) and were included in the analysis. Eighteen distinct topics were revealed from these, or 19, if the outliers are included. The outlier topic (topic -1) is assigned to 19.5% of the documents, which is within the expected range and reflects the heterogeneous nature of the exhibition content. A topic label is assigned to each topic in two steps. First, BERTopic provides a list of the most distinctive keywords and a keyword label for each topic (Table A.1). Based on these, an LLM is used to generate topic labels and descriptions (Table A.2). The thematic coherence revealed among the 18 topics indicates the need for additional thematic consolidation. The macro-topics are based on two criteria: the primary criterion is intertopic distance (Figure 4), while the secondary criterion is thematic coherence, which groups topics with semantic relatedness. Seven macro-topics are revealed and LLM topic labels and descriptions are added (Table 3). Macro-topics I, II, III, V and VI are based on intertopic distance, while macro-topics IV and VII are attributable to thematic coherence. Close proximity (<1.5) can be detected in macro-topics I, II and VI. This is followed by a moderate distance (1.5-2.0) in macro-topics III, IV and V, and the biggest distance (3.2) can be detected in macro-topic VII (>2.0). In addition, macro-topic VIII, *Cross-sector & adjacent topics*, includes topics at the edge of autonomous driving, such as plant-based materials for vehicle interiors and supply chain digitalisation. This topic is excluded from further analysis.

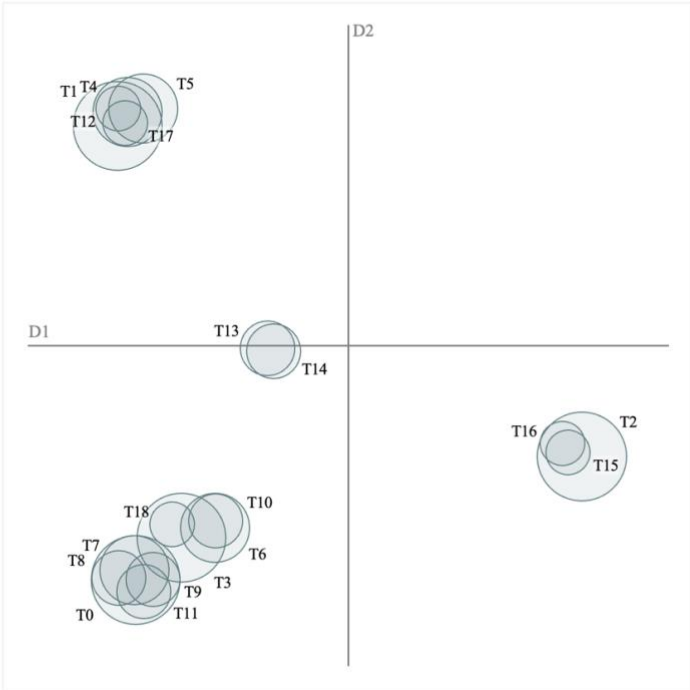

Figure 4 Intertopic distance map

Table 3 Macro-topic labels and descriptions

ID	(LLM) macro-topic label	Topics	d_{max}	n	(LLM) macro-topic description (short)
I	Camera systems & ADAS integration	13, 14	0.9	16	This cluster covers thermal cameras, driver monitoring systems and ADAS sensor integration technologies. Exhibitors demonstrate complete sensor suites with cost-effective integration tools for vehicle safety implementations.
II	Multi-modal sensor integration systems	2, 15, 16	1.1	37	This cluster covers radar, LiDAR and multi-camera sensor integration technologies for automotive applications. Exhibitors demonstrate 3D perception systems with multi-sensor fusion capabilities for autonomous vehicles.
III	Software-defined vehicle computing platforms	1, 4, 5	1.7	61	This cluster covers AI-powered software platforms, cloud services and high-performance computing systems for software-defined vehicles. Exhibitors demonstrate autonomous driving cockpits, vehicle update managers and automotive-grade computing platforms with sensor integration.

IV	ML training infrastructure & data pipelines	12, 7	1.5	21	This cluster covers cloud-based ML training platforms, data collection pipelines and GPU-accelerated deep learning workflows. Solutions demonstrate scalable ADAS development, raw data processing and reference model training for autonomous systems.
V	Urban mobility research & testing	3, 6, 18	1.6	42	This cluster covers Level 4 autonomous driving systems, digital mobility twins and urban transit technologies. Solutions demonstrate real passenger operations, demand-oriented transport models and citizen engagement platforms.
VI	Driver assistance & training systems	8, 11	1.2	19	This cluster covers hands-off driver assistance systems and interactive coaching technologies for driving behaviour improvement. Exhibitors demonstrate chauffeur-like navigation capabilities and immersive training platforms with real-time driver insights.
VII	Positioning & connected vehicle systems	0, 10	3.2	36	This cluster covers GNSS positioning, V2X communication and synchronised camera systems for connected vehicles. Exhibitors demonstrate teleoperated driving solutions and plug-and-play ADAS connectivity for remote control applications.
VIII	Cross-sector & adjacent topics	9, 17	-	15	-

In addition to the brief two-sentence descriptions (Table 3), the long LLM macro-topic descriptions are derived from the 18 topic keywords, labels and descriptions (Tables A.1 and A.2). These are, in turn, based on representative document samples from each topic. The long LLM descriptions are displayed below, providing information on the subject, technical details and applications, and trends and significance for each macro-topic.

I. Camera systems & ADAS integration

Camera Systems & ADAS (Advanced Driver Assistance Systems) Integration comprises 2 topics with 16 documents (6.5% of corpus), encompassing thermal camera and mirror systems alongside ADAS sensor integration and optimization technologies. This macro-cluster covers thermal imaging solutions that extend nighttime perception capabilities beyond traditional headlight range, while mirror-integrated camera systems provide comprehensive driver monitoring functionality. Exhibitors demonstrate complete ADAS sensor suites featuring analog sensor integration with advanced video link technologies and optimized connector solutions for enhanced bandwidth management. The cluster reveals industry emphasis on cost-effective safety implementations through integrated sensor architectures. This convergence of thermal imaging, driver monitoring, and *comprehensive ADAS integration represents critical infrastructure for autonomous vehicle perception systems, addressing both environmental sensing and occupant safety requirements essential for next-generation automotive applications.*

II. Multi-modal sensor integration systems

*Multi-Modal Sensor Integration Systems comprises 3 topics with 37 documents (15.0% of corpus), encompassing Radar (*Radio Detection And Ranging) *& LIDAR (*Light Detection And Ranging) *Sensor Integration, LIDAR Sensing Technology & Applications, and Multi-Camera & Radar Sensor Systems. This cluster addresses comprehensive sensor fusion technologies combining radar, LIDAR, and camera systems for automotive applications. Exhibitors demonstrate 3D perception systems with high-resolution measurement capabilities and optical components for autonomous vehicle navigation. Advanced LIDAR solutions feature low-profile designs with sequential flash scanning and multi-echo capabilities for enhanced object detection and classification. Multi-camera configurations integrate front and corner radar units, offering ranging capabilities for driver monitoring and cabin sensing applications. The convergence of these sensing modalities reflects the industry's shift toward redundant, complementary sensor architectures that enhance reliability and performance in diverse driving conditions. This integration approach is critical for achieving the robust environmental perception required for higher levels of vehicle autonomy, addressing safety-critical applications where single-sensor limitations could compromise system performance.*

III. Software-defined vehicle computing platforms

*Software-Defined Vehicle Computing Platforms comprises 3 topics with 61 documents (24.7% of corpus), integrating AI-powered human-machine interaction systems, cloud-based SDV (*Software-Defined Vehicle) *platforms, and high-performance computing architectures. The cluster encompasses diverse technologies including integrated display systems for automotive cockpits, collaborative development tools, and Intel Xeon CPU-based computing platforms with GPU acceleration. Exhibitors demonstrate vehicle update managers, data collection services, cloud analytics for connected vehicle functions, and sensor integration capabilities for autonomous driving applications. Industrial software platforms support cockpit development while automotive-grade SoC (*System-on-Chip) *solutions and SDK (*Software Development Kit) *frameworks enable high-performance processing requirements. These developments reflect the industry's transition toward software-centric vehicle architectures, where traditional hardware-defined systems evolve into flexible, updatable platforms. This transformation is critical for autonomous driving advancement, enabling real-time AI processing, seamless human-vehicle interaction, and continuous capability enhancement through over-the-air updates, positioning software-defined approaches as fundamental to next-generation automotive ecosystems.*

IV. ML training infrastructure & data pipelines

ML Training Infrastructure & Data Pipelines comprises 2 topics with 21 documents (8.5% of corpus), integrating autonomous vehicle data collection systems with cloud-based ML training platforms. The cluster encompasses comprehensive data generation workflows, from raw sensor data visualization to test scenario mining for reference model *development. Exhibitors demonstrate cloud-based deep learning platforms leveraging GPU infrastructure to accelerate ADAS development workflows, alongside specialized training studios for autonomous driving applications. Solutions span from data collection pipelines to scalable robotics training systems. This cluster reflects the industry's shift toward cloud-native ML infrastructure and automated data processing pipelines, addressing critical scalability challenges in autonomous vehicle development. The convergence of robust data collection methodologies with distributed training capabilities represents essential infrastructure for advancing commercial autonomous driving technology.*

V. Urban mobility research & testing

Urban Mobility Research & Testing comprises 3 topics with 42 documents (17.0% of corpus), integrating urban mobility research projects, Level 4 autonomous public transit systems, and navigation services. The cluster encompasses digital mobility twin systems that connect urban and rural regions through demand-oriented transport models, incorporating citizen engagement platforms for metropolitan testing environments. Exhibitors demonstrate Level 4 autonomous vehicles operating with real passengers on German roads through government-funded research initiatives, while shuttle services showcase hands-free driving capabilities and route optimization for daily commuting scenarios. Advanced AI-driven technologies enable public transportation systems to operate autonomously in complex urban environments, supported by comprehensive testing frameworks that validate safety and efficiency metrics. This research area represents a critical convergence of autonomous driving technology with public mobility infrastructure, addressing the growing demand for sustainable urban transportation solutions. The integration of citizen participation in testing phases indicates industry recognition that successful autonomous mobility deployment requires comprehensive stakeholder engagement and real-world validation beyond controlled environments.

VI. Driver assistance & training systems

Driver Assistance & Training Systems comprises 2 topics with 19 documents (7.7% of corpus), integrating hands-off driver assistance technologies and interactive training platforms. The cluster encompasses chauffeur-like navigation systems that handle standard driving functions across various road types, bridging manual and autonomous capabilities through eyes-hands coordination technologies. Exhibitors demonstrate immersive coaching platforms that provide real-time behavioural insights for driver improvement, while assisted driving solutions offer comprehensive all-scenario support. Interactive engagement systems focus on modifying driving habits through insurance-linked feedback mechanisms. These developments reflect the industry's dual approach of advancing automated assistance while simultaneously enhancing human driver competency. This convergence represents a critical transitional phase in autonomous driving adoption, where technology simultaneously replaces and augments human capabilities, ensuring safer roads through both automated systems and improved driver education.

VII. Positioning & connected vehicle systems

Positioning & Connected Vehicle Systems comprises 2 topics with 36 documents (14.6% of corpus), integrating autonomous vehicle operation with remote control capabilities *and V2X communication systems with ADAS integration. The cluster encompasses critical positioning technologies including GNSS satellite systems, LiDAR-based 3D mapping, and high-accuracy road mapping solutions for complex navigation scenarios. Exhibitors demonstrate teleoperated driving workstations that enable human operators to remotely assist autonomous vehicles in challenging situations such as construction zones, while others showcase plug-and-play V2X connectivity solutions featuring synchronized camera arrays and spatial processing capabilities. These developments reflect the industry's dual focus on precise vehicle localization and seamless vehicle-to-everything communication infrastructure. This convergence is essential for achieving reliable autonomous driving deployment, as it addresses both the technical challenges of accurate positioning in complex environments and the connectivity requirements for coordinated traffic management systems.*

4 Discussion and conclusion

Rapid advancements in automation technologies and artificial intelligence are transforming the transportation sector as we know it, promising benefits such as reduced carbon emissions, shorter travel times and an enhanced travel experience overall (Stead and Vaddadi, 2019; Bissell *et al.*, 2020). The expected worldwide increase in the availability of autonomous vehicles is enormous in the coming years (Alqahtani, 2025). These rapid developments are rendering current, time-consuming forecasting methods obsolete and calling for a new approach that enables stakeholders to make timely, informed decisions (Ebadi, Auger and Gauthier, 2022). This study therefore introduces an innovative methodology for technology forecasting and trend identification, drawing on insights from technology fairs where cutting-edge technologies are showcased. A deep learning-based topic modeling approach was used to analyse large quantities of unique visual data, ensuring a comprehensive data pool. Eight macro-topics were identified, and a detailed analysis was conducted for seven of these, with the exception of macro-topic VIII, *Cross-sector & adjacent topics*.

The size distribution of macro-topics reveals the strategic priorities within the AV technology development. The largest macro-topic, *III. Software-defined vehicle computing platforms* (24.7%), highlights the industry's transformation towards software-centric vehicle architecture. Cloud services, real-time AI processing on high-performance computing platforms and over-the-air updates are defining factors of competitive advantage, moving away from traditional hardware-defined systems. As emphasised by the findings of the descriptive analysis, the automotive industry is moving towards a service- and platform-based economy, in which companies compete based on ecosystems rather than one-off vehicle sales. The second largest macro-topic (17.0%) is macro-topic *V. Urban mobility research & testing*, which addresses the growing demand for sustainable, autonomous public transport solutions connecting urban and rural regions. Testing is carried out in metropolitan settings with real passengers to recognise the benefits of active stakeholder engagement and real-world validation beyond controlled environments. The third largest macro-topic is *II. Multi-modal sensor integration systems* (15.0%), which emphasises multi-modal architectures as the industry standard. Rather than relying on a single sensing modality, aspects such as radar, LiDAR and cameras are combined through sensor fusion to provide redundancy and complementary strengths.

This enables robust environmental perception throughout diverse conditions. Macro-topic *VII. Positioning & connected vehicle systems* (14.6%) is almost identical in size. It provides precise vehicle localisation and seamless vehicle-to-everything communication, which is critical infrastructure for coordinated traffic management. By integrating remote-control capabilities and vehicle connectivity, the macro-topic represents beyond-vehicle intelligence, where decision-making is either carried out through remote human control or distributed networked coordination. Although smaller in size, the remaining macro-topics are showing essential supporting infrastructure, especially for the transitional phase. Macro-topic *IV. ML training infrastructure & data pipelines* (8.5%) represents cloud-based training infrastructure and data processing pipelines that guarantee scalable, continuous learning loops. Generating simulated scenarios improves AV deployment by addressing long-tail edge cases that real-world data cannot capture. This is followed by macro-topic *VI. Driver assistance & training systems* (7.7%), which highlights the human component during the transition phase towards fully automated vehicles. The smallest macro-topic, *I. Camera systems & ADAS integration* (6.5%), addresses camera-centric sensing applications, including thermal imaging and ADAS camera integration. A comparison of the size of this topic with that of *multi-modal sensor integration systems* confirms the decline of camera-only strategies.

A coherent thematic analysis of MTs reveals a comprehensive autonomous driving system architecture, with the individual components having been addressed in prior research (Hakak *et al.*, 2022; Yang *et al.*, 2026; Fontan-Costas *et al.*, 2025). However, the findings go beyond previous work by revealing a four-layered structure that is consistent with the layered modular architecture associated with digital innovation (Yoo, Henfridsson and Lyytinen, 2010). The four layers are defined as follows: data acquisition; compute; intelligence; and application (Figure 5). The data acquisition layer comprises two data streams: on-board sensor data (from cameras, LiDAR and radar) and external positioning data (from GNSS, HD maps and V2X communication) [MTs I., II. and VII.]. Based on this layer, the compute layer fuses and processes the data to provide a coherent understanding of the environment, enabling real-time decision-making [MT III.]. Next, the intelligence layer uses cloud-based infrastructure to continuously trains and improve the models deployed on the compute layer [MT IV.]. Finally, user-facing services are operated in the application layer, providing value to the end user (e.g. urban mobility services) [MTs V. and VI.]. This alignment with the organising logic of digital innovation suggests that the automotive industry is undergoing a fundamental transformation towards a platform-based structure, similar to that of technology companies. Essentially, an autonomous vehicle will be a computing platform that cooperates with heterogeneous firms to provide novel components, such as software, cloud infrastructure and AI. This is similar to the way Google operates, simultaneously functioning as both a product and a platform across multiple layers (Yoo, Henfridsson and Lyytinen, 2010).

Our study has far-reaching implications for automotive producers and suppliers. It presents the latest technology trends in AV research and development, offering insights into the current state of autonomous transportation system adoption and its wider impact on societal dynamics. Consequently, this information provides a basis for forecasting and strategically planning the future development and widespread adoption of autonomous driving. Notably, however, it also provides context for disaster preparedness and the development of resilience strategies. Together, the thematic results revealed a four-layered modular architecture that is consistent with digital innovation, reflecting the industry's shift towards software-driven, service- and platform-based structures, rather than hardware-driven ones. This study has significant implications for research, as well as being relevant to industry practitioners and policymakers. Using AI-driven analytics with a unique visual dataset, the study introduces a methodology that enables practitioners and researchers to efficiently identify emerging technology trends and forecast advancements. By applying BERTopic to real-world technology fair data, the study presents an accessible approach to topic modeling and offers a streamlined, transferable workflow for future trend analysis across domains.

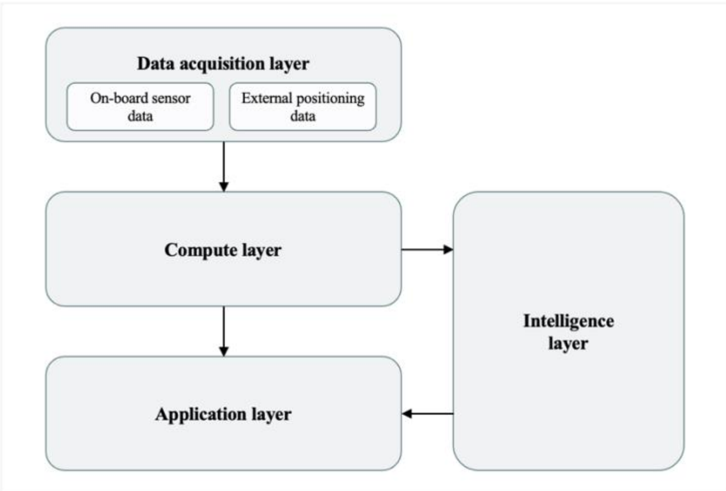

Figure 5 Four-layered modular architecture

While our study provides valuable insights within its scope, it is important to acknowledge its limitations. Since BERTopic relies on semantic similarity, it may fail to capture functional interdependencies between data with dissimilar terminology. Therefore, manual checking during macro-topic creation was necessary to ensure the most appropriate clustering for the purposes of this study. Another relevant aspect of the data is its focus mainly on the German automotive industry, given the location of the technology fair and participants. Although, also China and US participants were high in number and at all participants out of 20 countries give some proxy of a global representation of technology. Furthermore, the nature of a technology fair meant that critical components for the successful implementation of autonomous driving were not represented. These include cybersecurity, infrastructure, regulatory compliance and safety certification. Future work could address this by applying the presented methodology to fairs across different regions and specialised domains, including those in the fields of cybersecurity and regulation, to validate and extend its generalisability. Applying the methodology to successive fairs enables longitudinal, dynamic monitoring of industry and consumer trends and allows continuous flexible technology forecasting. Ultimately, we hope that scholars will build upon this work to refine the methodological approach, delivering timely, actionable insights and bridging the gap between science and industry while responsibly and thoughtfully leveraging AI-driven analytics.

Appendix A

Table A. 1 Topic keywords

ID	n	(Bertopic) keyword label	(Bertopic) top keywords
-1	60	Outliers	-
0	27	Road & 3d & Gnss	road, 3d, gnss, positioning, lidar, mapping, roads, satellite, accuracy, applications
1	26	Al & Human & Interaction	al, human, interaction, display, knowledge, industrial, support, integrated, research, requirements
2	23	Radar & Range & Lidar	radar, range, lidar, optical, resolution, beam, measurement, field, 3d, lighting
3	21	Institut & Öffentlichen & Projekt	institut, öffentlichen, projekt, test, mobility, bed, lower, forschung, mobilität, demand
4	19	Software & Ai & Sdv	software, ai, sdv, cloud, defined, software defined, quality, services, platform, al
5	16	Chip & Computing & Soc	chip, computing, soc, grade, platform, bit, sdk, host, chips, ads
6	15	Stadt & Zukunft & Mobilität	stadt, zukunft, mobilität, raum, projekt, forschung, ki, verkehr, technologie, uo
7	13	Cloud & Ml & Learn	cloud, ml, learn, training, deep learning, adas, gpus, learning, scale, robots
8	10	Hands & Eyes & Description	hands, eyes, description, unique, eyes hands, assist, chauffeur, tech, odd, manual
9	9	Plant & Carbon & Interior	plant, carbon, interior, concept, shuttle, public, materials, public transport, passengers, transport
10	9	V2x & Os & Cameras	v2x, os, cameras, contact, synchronised, spatial, wn, plug play, adas, connection
11	9	Coaching & Insurance & Drivers	coaching, insurance, drivers, habits, improve, assisted, experience, insights, behaviour, engagement
12	8	Raw & Visualization & Reference	raw, visualization, reference, training, learning, studio, net, collects, test, example
13	8	Camera & Mirror & View	camera, mirror, view, temperature, temperature range, details, side, video, 40, voltage
14	8	Camera & Link & Video	camera, link, video, analog, connectors, bandwidth, sd, cost, higher, adas
15	7	Selection & Frame & Classification	selection, frame, classification, collection, predictions, labeling, edge, collected, search, profile
16	7	Passive & Camera & Offer	passive, camera, offer, cameras radar, measuring, front, problem, camera, radar, ranging
17	6	Group & Industrial & Jahren	group, industrial, jahren, euro, mobility, unternehmen, ceo, supply chain, gen, b3
18	6	Urban & Mobility & Mobilität	urban, mobility, mobilität, passengers, home, tops, 200, operations, shuttle, öffentlichen

Table A. 2 Topic labels and descriptions

ID	n	(LLM) topic label	(LLM) topic description
-1	60	Outliers	Documents that did not fit any identified topic cluster with sufficient confidence.
0	27	Autonomous Vehicle Operation & Remote Control	This cluster addresses teleoperated driving systems for autonomous vehicles requiring human operator assistance. Applications include remote control workstations for handling complex road situations like construction sites.
1	26	AI-Powered Automotive Software & Human-Machine Interaction	This cluster addresses AI-powered human-machine interaction systems with integrated display technologies for automotive applications. Companies demonstrate industrial software platforms supporting autonomous driving cockpits and collaborative development tools meeting automotive industry requirements.
2	23	Radar & LIDAR Sensor Integration	This cluster covers radar and lidar sensors with optical components for automotive applications. Exhibitors demonstrate 3d perception systems, lighting integration panels and high-resolution measurement capabilities for autonomous vehicles.
3	21	Urban Mobility Research & Testing Projects	This cluster addresses autonomous mobility project development connecting urban and rural regions through demand-oriented transport models. The documents show digital mobility twin systems and citizen engagement platforms for testing metropolitan mobility solutions.
4	19	Software-Defined Vehicle Cloud Platform	This cluster addresses SDV cloud platform services for software-defined vehicle ecosystems. Exhibitors demonstrate vehicle update managers, data collectors and cloud-based analytics services for connected vehicle functions.
5	16	High-Performance Computing Platforms for Autonomous Vehicles	This cluster covers automotive-grade computing platforms with Intel Xeon CPUs and GPU chips. Exhibitors showcase high-performance computing systems for autonomous vehicle applications with sensor integration capabilities.
6	15	Autonomous Public Transit Level 4 Systems	This cluster addresses Level 4 autonomous driving technology for public transport systems in urban environments. The project demonstrates autonomous vehicles operating with real passengers on German roads through government-funded research initiatives.
7	13	Cloud-Based ML Training for Autonomous Systems	This cluster covers cloud-based ML training platforms for ADAS development and scalable autonomy. Solutions accelerate deep learning workflows for autonomous driving technology using cloud infrastructure and GPUs.
8	10	Hands-Off Driver Assistance Systems	This cluster addresses hands-off driver assist systems bridging manual and autonomous driving capabilities. Products offer chauffeur-like navigation tech handling standard driving functions across various road types.

9	9	Plant-Based Interior Materials for Public Transport	This cluster covers plant-based interior materials for public transport vehicles and shuttle concepts. Exhibitors showcase carbon-conscious materials like plant-based polyester for passenger seating and interior surfaces.
10	9	V2X Communication & ADAS Integration	This cluster addresses V2X communication systems and ADAS integration for connected vehicles. Exhibitors demonstrate plug-and-play connectivity solutions with synchronised camera systems and spatial processing capabilities.
11	9	Interactive Driver Training & Assisted Driving Systems	This cluster addresses assisted driving coaching systems that provide interactive engagement for driver behaviour improvement. Exhibitors demonstrate immersive experience platforms and all-scenario assisted driving solutions with real-time insights.
12	8	Autonomous Vehicle Data Collection & Training Pipeline	This cluster addresses data collection and training systems for autonomous driving development. Solutions include raw data generation, test scenario mining and reference model training workflows.
13	8	Thermal Camera & Mirror Systems	This cluster addresses thermal camera systems and driver monitoring camera technology for automotive applications. Products include thermal imaging for nighttime perception beyond headlight range and mirror-integrated camera solutions for driver monitoring.
14	8	ADAS Sensor Integration & Optimization	This cluster addresses ADAS camera and analogue sensor integration systems for automotive applications. Exhibitors demonstrate complete ADAS sensor suites with integration tools for cost-effective vehicle safety implementations.
15	7	LIDAR Sensing Technology & Applications	This cluster covers high-resolution LiDAR sensors with perception software for object detection and classification. Products feature low-profile designs with sequential flash scanning and multi-echo capabilities for automotive applications.
16	7	Multi-Camera & Radar Sensor Systems	This cluster addresses multi camera systems with front and corner radars for automotive sensing applications. Exhibitors offer camera heads with radar ranging capabilities for driver monitoring and cabin sensing.
17	6	Industrial Digital Transformation & Supply Chain	This cluster addresses industrial group mobility transformations using digital supply chain ecosystems. The documents show metaverse implementations generating euro savings and automotive software development for major companies.
18	6	Urban Mobility & Navigation Services	This cluster addresses urban mobility services and public transportation operations for passengers. Solutions include shuttle services, hands-free driving capabilities and route optimisation systems for daily commuting scenarios.

References and Notes

Allaoui, M., Kherfi, M. L. and Cheriet, A. (2020) 'Considerably Improving Clustering Algorithms Using UMAP Dimensionality Reduction Technique: A Comparative Study', pp. 317-325.

Alqahtani, T. (2025) 'Recent Trends in the Public Acceptance of Autonomous Vehicles: A Review', *Vehicles,* 7, pp. 45.

Bissell, D., Birtchnell, T., Elliott, A. and Hsu, E. L. (2020) 'Autonomous automobilities: The social impacts of driverless vehicles', *Current Sociology,* 68(1), pp. 116-134.

Blei, D. M. (2012) 'Probabilistic topic models', *Commun. ACM,* 55(4), pp. 77–84.

Campello, R., Moulavi, D. and Sander, J. (2013) *Density-Based Clustering Based on Hierarchical Density Estimates.*

Cohen, T., Stilgoe, J., Stares, S., Akyelken, N., Cavoli, C., Day, J., Dickinson, J., Fors, V., Hopkins, D., Lyons, G., Marres, N., Newman, J., Reardon, L., Sipe, N., Tennant, C., Wadud, Z. and Wigley, E. (2020) 'A constructive role for social science in the development of automated vehicles', *Transportation Research Interdisciplinary Perspectives,* 6, pp. 100133.

Cumming, D. J. and Johan, S. A. (2013) *Venture Capital and Private Equity Contracting: An International Perspective. Elsevier insights*: Elsevier.

Daim, T. U., Rueda, G., Martin, H. and Gerdsri, P. (2006) 'Forecasting emerging technologies: Use of bibliometrics and patent analysis', *Technological Forecasting and Social Change,* 73(8), pp. 981-1012.

Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2018) *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.*

Ebadi, A., Auger, A. and Gauthier, Y. (2022) *Detecting Emerging Technologies and their Evolution using Deep Learning and Weak Signal Analysis.*

Eurostat (2025) *NACE Rev. 2.1 – Statistical classification of economic activities in the European Union – 2025 edition*. Available at: https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/w/ks-gq-24-007.

Fontan-Costas, B., Diaz-Cacho, M., Fernandez-Boullon, R., Alonso-Carracedo, M. and Perez-Robles, J. (2025) *A Modular Architecture Design for Autonomous Driving Racing in Controlled Environments.*

Grootendorst, M. (2022) 'BERTopic: Neural topic modeling with a class-based TF-IDF procedure'.

Hakak, S., G*adekallu, T.*, *P*riya, S., M, P., Reddy, P., De Alwis, C. and Liyanage, M. (2022) *Autonomous Vehicles in 5G and Beyond: A Survey.*

Mattas, K., Botzoris, G. and Papadopoulos, B. (2021) 'Safety aware fuzzy longitudinal controller for automated vehicles', *Journal of Traffic and Transportation Engineering (English Edition),* 8(4), pp. 568-581.

McInnes, L., Healy, J. and Astels, S. (2017) 'hdbscan: Hierarchical density based clustering', *The Journal of Open Source Software,* 2, pp. 205.

McInnes, L., Healy, J., Saul, N. and Grossberger, L. (2018) 'UMAP: Uniform Manifold Approximation and Projection', *Journal of Open Source Software,* 3, pp. 861.

Metrick, A. and Yasuda, A. (2021) *Venture Capital and the Finance of Innovation.* Wiley.

Reimers, N. and Gurevych, I. (2019) *Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks.*

Reis, J., Muñoz Tabora, J., Lima, M., Monteiro, F., Monteiro, S., Bezerra, U. and Tostes, M. (2025) 'Medium and Long Term Energy Forecasting Methods: A Literature Review', *IEEE Access,* PP, pp. 1-1.

Stead, D. and Vaddadi, B. (2019) 'Automated vehicles and how they may affect urban form: A review of recent scenario studies', *Cities,* 92, pp. 125-133.

Tsai, P.-H., Berleant, D., Segall, R., Hyacinthe, A., Batthula, V., Batthula, R., Duggirala, S. and Howell, M. (2023) *Quantitative Technology Forecasting: a Review of Trend Extrapolation Methods.*

Verband der Automobilindustrie e. V. (2025) *IAA Mobility 2025*. Available at: https://www.iaa-mobility.com/en (Accessed: 09.09.2025.

Yang, L., Cheng, J., Zhou, M., Liu, C., Ni, Z., Xie, M. and Gao, S. (2026) 'Integrated Perception, Communication, and Computation for Autonomous Vehicle and Road Infrastructure Network', *IEEE Transactions on Mobile Computing*.

Yao, Q., Tian, Y., Wang, Q. and Wang, S. (2020) 'Control Strategies on Path Tracking for Autonomous Vehicle: State of the Art and Future Challenges', *IEEE Access,* 8, pp. 161211-161222.

Yoo, Y., Henfridsson, O. and Lyytinen, K. (2010) 'Research Commentary: The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research', *Information Systems Research,* 21(4), pp. 724-735.