CogSust

Cognitive Sustainability

Mar 2026 Vol. 5. No. 1

ISSN 2939 - 5240

Cognitive Sustainability

Cognitive Sustainability (CogSust) is a double-blind peer-reviewed scientific journal published by CogSust Ltd.

(H1116 Budapest Putnok u 9.)

The person responsible for publishing: Mária Szalmáné Csete editor@cogsust.com

The person responsible for editing: Ádám Török info@cogsust.com

CogSust is an online quarterly journal, publication frequency: quarterly, by

March, June, September, December.

ISSN 2939-5240

The journal is indexed by:

Library of Hungarian Scientific Works

Repository of the Library of Hungarian Academy of Science

Google Scholar

Sherpa Romeo

RePEc (Research Papers in Economics)

Directory of Open Access Journals

Crossref

Semantic Scholar

ERIH PLUS

Index Copernicus International

Dimensions

Scilit

Lens

https://doi.org/10.55343/CogSust.21280

Road transport safety and children's cognitive attitudes

Viktoria Otvos
Department of Transport Technology and Economics;
Faculty of Transport Engineering and Vehicle Engineering;
Budapest University of Technology and Economics
Budapest, Hungary
otvos.viktoria@edu.bme.hu

Abstract
This paper summarizes evidence on road transport safety with a focus on children's cognitive attitudes and pedestrian behaviour. It integrates findings on attention, visual processing, executive functions, risk perception, knowledge–behaviour transfer, and environmental complexity. Consistent evidence indicates that developmental limitations in attention and processing speed constrain safe crossing decisions in younger children. Knowledge-focused education alone does not reliably improve real-world behaviour, and behavioural training targeting procedural skills yields modest safety improvements. Built environment features – particularly traffic speed and volume, as well as visual clutter – systematically shape both perceived and objective safety outcomes. The paper concludes with implications for training, urban design, and family practices, and outlines directions for future research.

Keywords
road transport safety, child pedestrian behaviour, cognitive development, risk perception, behavioural training

1. Introduction

Road transport injuries remain a leading cause of death and disability among children worldwide. Within this public health challenge, child pedestrians represent a particularly vulnerable group because safe participation in traffic requires integrated cognitive, perceptual, and motor skills that are still developing during early and middle childhood (Vijay et al., 2024; Zeedyk et al., 2001). Globally, road traffic injuries remain the leading cause of death for children and young people aged 5–29 years, with an estimated 1.19 million fatalities annually. More than half of these fatalities occur among vulnerable road users – pedestrians, cyclists, and motorcyclists – especially in low- and middle-income countries (WHO, 2023).

Beyond mortality, the morbidity and disability burden impose substantial socioeconomic costs on families and health systems. The United Nations' Global Plan for the Decade of Action for Road Safety 2021–2030 (WHO, 2021) targets a 50% reduction in road traffic deaths and injuries by 2030. It explicitly frames child safety within Safe System principles: human life is paramount; predictable errors must not be fatal; responsibility is shared across system actors. Children's active travel confers major developmental, health, and environmental benefits but exposes them to complex traffic environments. Systematic evidence indicates that high traffic speeds and volumes consistently reduce both objective and perceived safety for children, whereas sidewalk presence and protected crossings improve safety (Amiour et al., 2022; Cloutier et al., 2021).

Environmental visual clutter – e.g., billboards – adds extraneous cognitive load, increasing missed crossing opportunities and widening gaze dispersion in children compared with adults (Tapiro et al., 2020). Developmental research indicates that inflexion points in pedestrian competence occur: a strategic shift in visual search emerges around ages 7–8, and general competence is typically achieved by about age 10 (Whitebread and Neilson, 2000). Younger children rely on simpler heuristics (e.g., distance gap) and are markedly more likely to make unsafe crossing decisions. Cognitive functioning subsumes chronological age as a predictor of risky route selection (Oxley et al., 2007; Barton et al. 2012; Tabibi and Pfeffer, 2003). Crucially, classroom education successfully increases declarative road-safety knowledge with months-long retention, yet it does not translate into improved real-world crossing behaviour. In contrast, behaviourally oriented training that targets procedural competencies yields small-to-moderate improvements sustained for months (Zeedyk et al., 2001; Schwebel et al., 2014). Protecting children in traffic, therefore, requires

https://doi.org/10.55343/CogSust.21280

capability building under realistic cognitive load, complemented by Safe System design that reduces the demands placed on developing cognitive systems. Finally, recent syntheses highlight the underrepresentation of LMICs (low- and middle-income countries) and the need for exposure-adjusted injury data linked to intervention evaluations (Vijay et al., 2024; Wazana et al., 1997). Safeguarding children in road transport is not only a scientific goal and a policy imperative, but also aligns with the SDGs (Sustainable Development Goals) on health, equity, and sustainable cities.

In this paper, it is hypothesised that children's pedestrian safety is primarily limited by developmental constraints in attention, visual search efficiency, and processing speed. A literature review is conducted to test the hypothesis and find effective ways of intervention. The rest of the article is structured as follows: the Methodology section describes how evidence from empirical and review studies was extracted. Results summarise key findings across cognitive domains, behaviours, interventions, and environmental/social moderators with child pedestrians. The Analysis section integrates mechanisms that reconcile apparent contradictions (knowledge–behaviour gap), and the Discussion outlines implications for policy, urban design, education, and family practices. The Conclusion assesses the hypothesis and proposes future research directions.

2. Methodology

A detailed bibliographic search on sciencedirect.com was conducted. Sources included empirical studies, systematic reviews, and meta-analyses on child pedestrian safety, cognitive factors, and built-environment influences. Study characteristics, cognitive constructs (attention, visual processing, executive functions, risk perception), behavioural outcomes (crossing judgments, route selection, dash-out), intervention modalities (education vs. behavioural training), and environmental/social moderators (traffic speed/volume, visual clutter, sidewalks, parenting styles) were extracted.

https://doi.org/10.55343/CogSust.21280

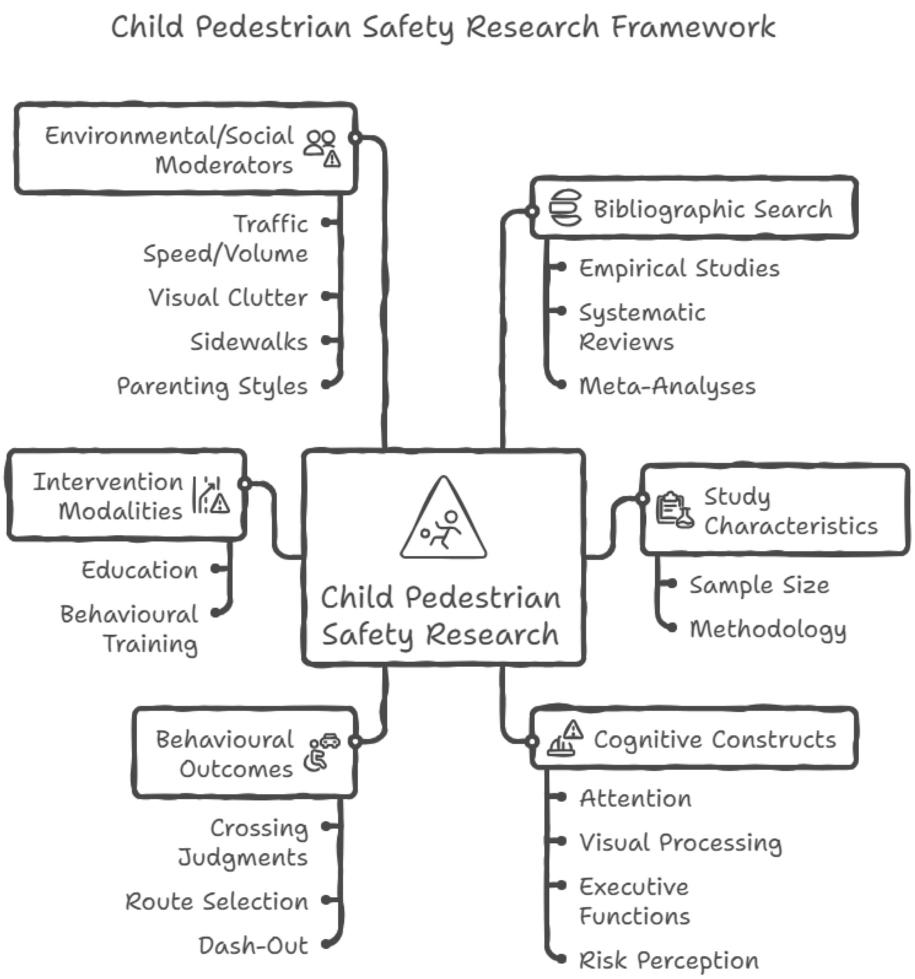

Figure 1. Child Pedestrian Safety Research framework
Source: own compilation with napkin.ai

3. Results

Children's ability to identify safe crossing sites and resist attentional interference improves with age, with a strategic shift around 7–8 years and general competence by ~10 years. Lower perceptual, attentional, and executive performance predicts unsafe crossing decisions; cognitive functioning can subsume age as a predictor of risky route selection (Amiour et al., 2022). Educational interventions reliably increase declarative knowledge about safe/dangerous crossing locations, with retention lasting up to six months; however, they do not improve real-world crossing behaviour compared with controls (Tapiro et al., 2020). Individualised/small-group training targeting dash-out prevention, crossing at parked cars, and safe route selection produces small to moderate safety gains, sustained for 2–8 months (Oxley et al., 2007). High vehicle speed/volume consistently associates with unsafe perceptions and increased injuries, while sidewalk presence improves perceived and objective safety. Visual clutter increases missed crossing opportunities and widens children's gaze dispersion relative to adults. Risk perception relates to age, gender, socioeconomic status, and parenting styles. Negative parenting (e.g., poor monitoring) associates with poorer risk perception

https://doi.org/10.55343/CogSust.21280

(Barton et al., 2012).

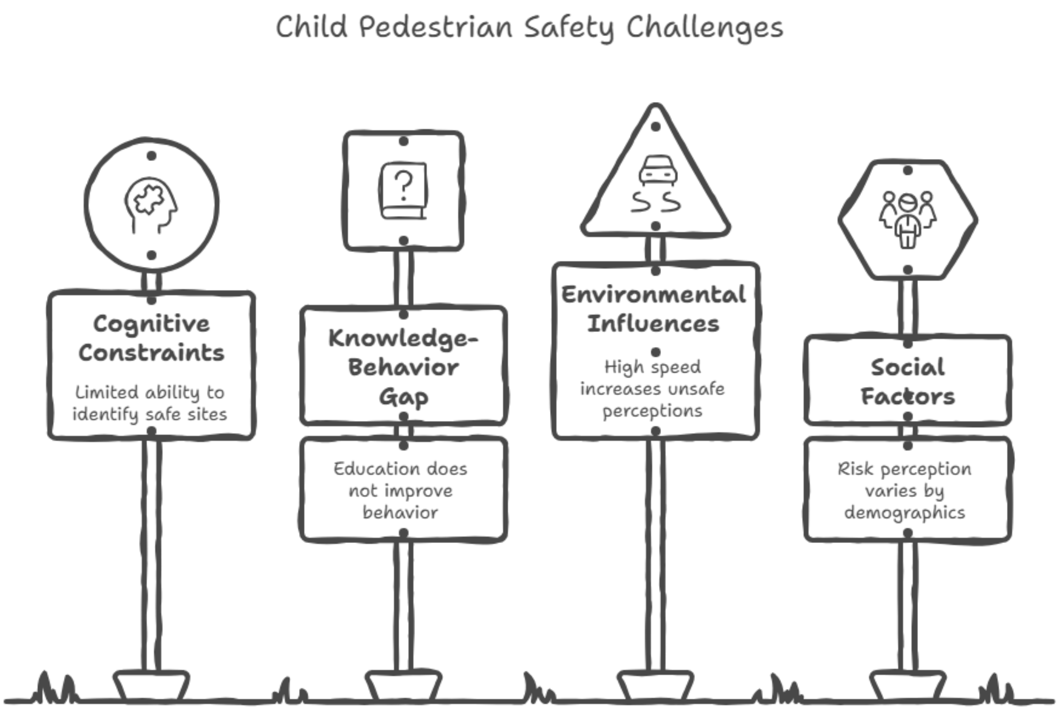

Figure 2. Child Pedestrian Safety challenges
Source: own compilation with napkin.ai

4. Analysis

Dual-process decision models help explain errors in child road crossing. Under time pressure, younger children rely on fast heuristics rather than effortful integration, which is limited by working memory and slower processing, and is biased towards distance-gap judgments, increasing error susceptibility when traffic streams are complex (Whitebread and Neilson, 2000; Oxley et al., 2007). Cognitive load theory clarifies the knowledge–behaviour gap: declarative schemas gained via classroom learning are not readily applied under high extraneous load (noise, visual clutter, multi-stream traffic), whereas procedural training automates gaze patterns and hazard detection, enabling transfer to real contexts (Zeedyk et al., 2001; Schwebel et al., 2014). VR (virtual reality) laboratory evidence shows that higher visual load reduces crossing opportunities and broadens gaze dispersion among children, consistent with immature inhibitory control (Tapiro et al., 2020). Population-level milestones (strategy shift at 7–8; competence by ~10) coexist with substantial within-age variability in executive functions (attention, inhibition, processing speed). Cognitive functioning can outperform age in predicting risky route selection, supporting screening for cognitive readiness and tailoring instructional intensity (Barton et al., 2012; Tabibi and Pfeffer, 2003).

https://doi.org/10.55343/CogSust.21280

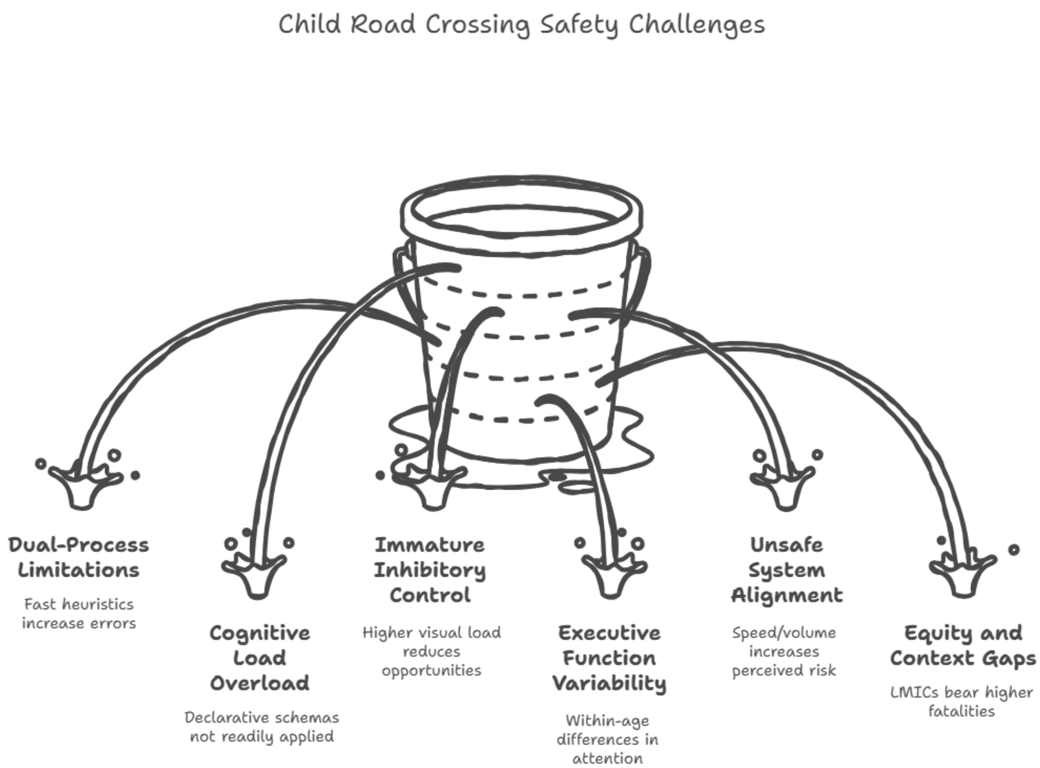

Figure 3. Child Pedestrian Safety challenges
Source: own compilation with napkin.ai

Speed management, protected crossings, and sidewalk continuity reduce cognitive demands and error consequences for child pedestrians. Reviews consistently identify speed/volume as primary determinants of both objective and perceived safety. The WHO's Global Plan centres on systemic responsibility to prevent typical human errors from being fatal (Amiour et al., 2022; Cloutier et al., 2021; WHO, 2021, 2023). LMICs bear a higher share of vulnerable user fatalities, as exposure patterns (longer walking distances), enforcement gaps, and infrastructure deficits elevate risk even for well-trained children. Social determinants – such as parental supervision and socioeconomic status – shape risk perception and behaviour, signalling the need for integrated family/community components (WHO, 2023; Wazana et al., 1997; Amiour et al., 2022). Education is necessary but insufficient without procedural fluency. Although age predicts average gains, executive functions predict individual outcomes, and environmental complexity actively modulates cognitive demand. A combined pathway – procedural training, exposure-aware urban design, and family/school practices – should be evaluated using robust, exposure-adjusted outcomes across diverse geographies (Vijay et al., 2024; Cloutier et al., 2021).

5. Discussion

The above results suggest that in education hands-on practice in realistic scenarios (on-street supervision and VR/AR) are to be prioritized to automate gaze strategies, hazard detection, and speed–distance judgments. Furthermore, curricula must be differentiated by developmental stage. In practice, this means that heuristics should be simplified for younger children and anticipatory strategies may be incorporated in the curriculum for older cohorts.

https://doi.org/10.55343/CogSust.21280

As far as urban design and policy are concerned, it is advisable to implement 30–40 km/h limits in school zones and residential streets. Moreover, raised crossings, curb extensions, protected mid-block crossings, and continuous sidewalks might also increase safety levels. The regulation of visual clutter near crossings may reduce extraneous load. Crossing-guard deployments should be aligned with risk-based site selection and monitored for behavioural compensation.

Parental modelling of safe road-crossing behaviour together with consistent adult supervision should be actively promoted as a core component of child pedestrian safety interventions. Socioeconomic barriers to safe transport can be mitigated through structural programmes such as Safe Routes to School and through investment into targeted crossing infrastructure.

Future research should systematically screen for children's cognitive readiness for independent road use, and should report exposure-adjusted outcomes (e.g. injuries per km walked; missed crossing opportunities; safe route selection rates). Embedding randomised or quasi-experimental comparisons of procedural versus knowledge-only curricula into future studies and linking outcomes to police/injury registries would also be promising. Studies in LMICs would also be advisable, as data remain sparse from those regions, while risks are the highest.

Regarding methodologies, eye-tracking studies should be carried out in realistic environments. Methods should be found to quantify the benefits of visual clutter reduction. Scalable digital training tools should be developed, together with an effective testing system for their efficiency. Finally, integrating Safe System analytics (e.g. kinetic energy thresholds, forgiving street design) with developmental metrics would enable the derivation of design standards optimised for child pedestrians.

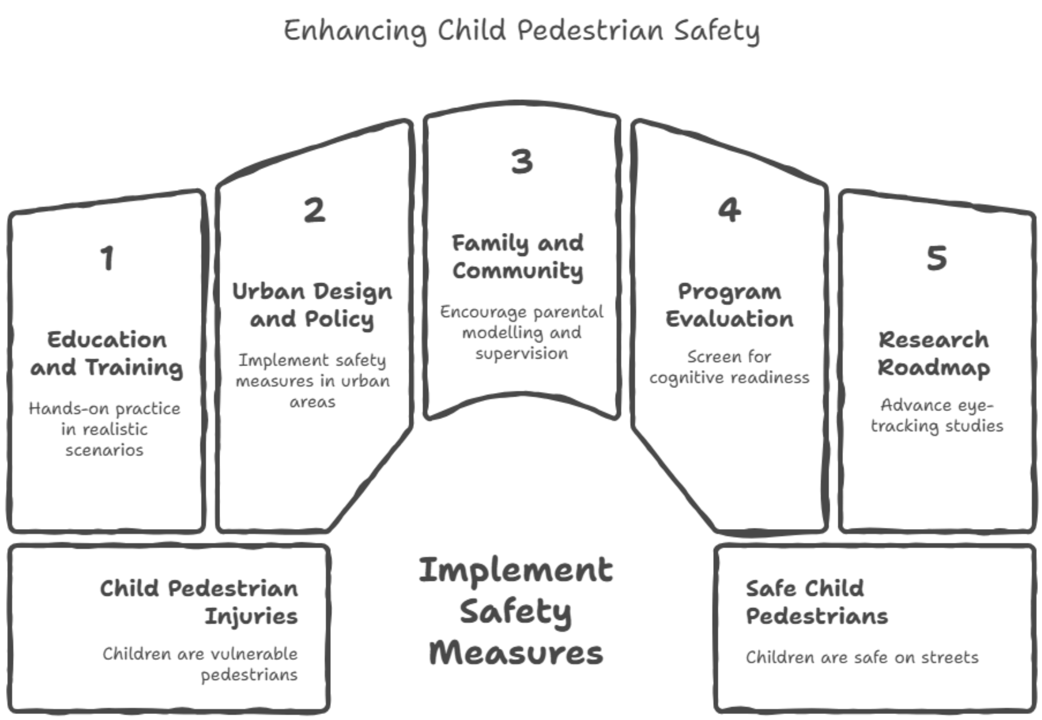

Figure 4. Child Pedestrian Safety challenges
Source: own compilation with napkin.ai

6. Conclusion

https://doi.org/10.55343/CogSust.21280

The evidence supports the hypothesis: developmental constraints in attention, visual search, and processing speed are primary limits on younger children's pedestrian safety. This means, interventions that train procedural skills under realistic cognitive load outperform declarative education in improving behaviour. Future research should design scalable behavioural training calibrated to cognitive readiness, quantify the benefits of speed management and visual clutter reduction near crossings, integrate parent-focused components, and extend evaluations to LMIC contexts to strengthen generalisability.

7. Acknowledgement

Project no. KDP-IKT-2023-900-I1-00000957/0000003 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary, financed under the National Research, Development and Innovation Fund, as part of the KDP-2023 funding scheme.

References

Amiour, Y., Waygood, E. O. D., van den Berg, P. E. W. (2022). Objective and perceived traffic safety for children: A systematic literature review of traffic and built environment characteristics related to safe travel. *International Journal of Environmental Research and Public Health*. 19(5), 2641. DOI: https://doi.org/10.3390/ijerph19052641

Barton, B. K., Ulrich, T., Lyday, B. (2012). The roles of gender, age and cognitive development in children's pedestrian route selection. *Child: Care, Health and Development*, 38(2), 280–286. DOI: https://doi.org/10.1111/j.1365-2214.2010.01202.x

Cloutier, M.-S., Beaulieu, E., Fridman, L., Macpherson, A. K., Hagel, B. E., Howard, A. W., Churchill, T., Fuselli, P., Macarthur, C., Rothman, L. (2021). State-of-the-art review: Preventing child and youth pedestrian motor vehicle collisions: Critical issues and future directions. *Injury Prevention*, 27(1), 77–84. https://doi.org/10.1136/injuryprev-2020-043829

Oxley, J. A., Congiu, M., Whelan, M. I., D'Elia, A. D., Charlton, J. L. (2007). The impacts of functional performance, behaviour and traffic exposure on road-crossing judgements of young children. *Annual Proceedings of the Association for the Advancement of Automotive Medicine*. 51, 81–96. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC3217496/

Schwebel, D. C., Barton, B. K., Shen, J., Wells, H. L., Bogar, A. L., Heath, G., & McCullough, D. (2014). Systematic review and meta-analysis of behavioral interventions to improve child pedestrian safety. *Journal of Pediatric Psychology*. 39(8), 826–845. DOI: https://doi.org/10.1093/jpepsy/jsu024

Tabibi, Z., Pfeffer, K. (2003). Choosing a safe place to cross the road: The relationship between attention and identification of safe and dangerous road-crossing sites. *Child: Care, Health and Development*. 29(4), 237–244. DOI: https://doi.org/10.1046/j.1365-2214.2003.00336.x

Tapiro, H., Oron-Gilad, T., Parmet, Y. (2020). Pedestrian distraction: The effects of road environment complexity and age on pedestrian's visual attention and crossing behavior. *Journal of Safety Research*. 72, 101–109. DOI: https://doi.org/10.1016/j.jsr.2019.12.003

Vijay, N. C., Rokade, S., Bivina, G. R. (2024). Towards safer streets: A review of child pedestrian behavior and safety worldwide. *Transportation Research Part F: Traffic Psychology and Behaviour*. 103, 638–660. DOI: https://doi.org/10.1016/j.trf.2024.05.005

Whitebread, D., Neilson, K. (2000). The contribution of visual search strategies to the development of pedestrian skills by 4–11 year-old children. *British Journal of Educational Psychology*. 70(4), 539–557. DOI: https://doi.org/10.1348/000709900158290

Zeedyk, M. S., Wallace, L., Carcary, B., Jones, K., & Larter, K. (2001). Children and road safety: Increasing knowledge does not improve behaviour. *British Journal of Educational Psychology*. 71(4), 573–594. DOI: https://doi.org/10.1348/000709901158686

Wazana, A., Krueger, P., Raina, P., Chambers, L. (1997). A review of risk factors for child pedestrian injuries: Are they modifiable? *Injury Prevention*. 3(4), 295–304. DOI: https://doi.org/10.1136/ip.3.4.295

WHO – World Health Organization (2021). *Global Plan for the Decade of Action for Road Safety 2021–2030*. Geneva: WHO. URL: https://www.who.int/publications/m/item/global-plan-for-the-decade-of-action-for-road-safety-2021-2030

WHO – World Health Organization (2023). *Global status report on road safety 2023*. Geneva: WHO. URL: https://digitallibrary.un.org/record/4030268?v=pdf&ln=hu

https://doi.org/10.55343/CogSust.22154

Safety management systems in road transport: Opportunities and challenges in shared mobility

Hazem Al-Mahamid a*
https://orcid.org/0009-0003-3073-3272
Karak Governorate Public Works Directorate, Roads Administration Directorate,
Amman, Jordan
hazem.alx1992@yahoo.com

Abstract

Safety Management Systems (SMS) have long been applied in high-risk sectors such as aviation and rail, but their adaptation to road transport remains underexplored. This paper examines the potential of SMS as a proactive framework for improving safety in road transport, with particular attention to shared vehicles and car sharing. It traces the origins and principles of SMS, evaluates its benefits and limitations, and analyses case studies from public transit, shared mobility platforms, and international experiences. The findings show that SMS can strengthen safety culture, enhance hazard identification, and support continuous monitoring, yet challenges such as fragmented governance, inconsistent driver practices, and technological barriers limit its effectiveness. The paper argues that stronger policy commitment, regulatory oversight, and institutional coordination are essential for successful implementation, alongside platform-level measures such as driver training, fatigue management, and data-driven monitoring. The distinctive contribution of this study lies in extending SMS from traditional high-risk sectors into the fragmented, rapidly evolving context of road transport and shared mobility, clarifying both the opportunities and the systemic challenges of such adaptation for policymakers, regulators, and platform operators.

Keywords: Safety Management Systems (SMS), Road transport safety, Shared vehicles, Car sharing, Policy and decision-making, Risk management, Digital mobility platforms.

1. Introduction

Over the past few decades, safety has become a central concern in road transport, not only for policymakers but also for operators and users. Mauriño (2017) explains that a Safety Management System (SMS) should not be seen as a simple regulatory requirement. Instead, it represents a comprehensive managerial approach that embeds safety into all organisational processes. This includes identifying hazards, assessing risks, developing safety policies, providing staff training, and continuously evaluating safety performance (Al-Mahamid et al., 2025; Mauriño, 2017).

The OECD/ITF further stresses that the successful implementation of SMS in the road transport sector depends on strong institutional commitment, clear policies, sufficient financial resources, and the active involvement of stakeholders, including governments, service providers, and road users. When these conditions are met, safety is no longer treated as a costly burden but as an added value that improves overall performance and reduces human and economic losses from accidents (OECD/ITF, 2018). Despite continuous progress in vehicle design and infrastructure, road traffic crashes remain one of the world's most pressing safety challenges. Each year, an estimated 1.19 million people lose their lives, and up to 50 million are injured, with vulnerable road users such as pedestrians, cyclists, and motorcyclists representing more than half of the victims. Beyond the human suffering, crashes impose a heavy economic burden, costing countries approximately 3% of their GDP annually (WHO, 2023). Recognising this challenge, the United Nations has included in its Sustainable Development Goals the ambitious target of reducing global road traffic deaths and injuries by 50% by 2030. (United Nations, 2015)

https://doi.org/10.55343/CogSust.22154

The roots of Safety Management Systems (SMS) go back to the aviation sector, where the International Civil Aviation Organisation (ICAO) began developing the concept in the late 1990s and early 2000s, particularly with the release of the Global Aviation Safety Plan (ICAO, 2013) and the adoption of a proactive approach to risk management. This framework was later adopted by other high-risk sectors, including railways and maritime transport, demonstrating its flexibility and effectiveness in reducing accidents and improving safety performance. (Mauriño, 2017).

In recent years, there has been a growing interest in applying Safety Management System (SMS) principles to the road transport sector. The increasing complexity of road environments drives this shift, as does the diversity of road users and persistently high accident rates. Agencies such as the U.S. Federal Transit Administration (FTA) have launched pilot programs to implement SMS in public transportation systems, including projects with the Chicago Transit Authority (CTA) and Maryland's Ride On bus service, aiming to enhance proactive safety performance and risk mitigation (FTA, 2020). Additionally, the National Transportation Safety Board (NTSB) has recommended adopting SMS across all modes of transportation, including road transport, emphasising its value in identifying hazards early and implementing preventive measures – regardless of the size or scope of the organisation (NTSB, 2023). These initiatives reflect a broader shift toward systematic and anticipatory safety management, moving beyond reactive approaches and fostering a stronger safety culture on the roads.

In the context of road transport governance, several studies emphasise that road safety can be managed through structured institutional systems and measurable strategic objectives. For example, Čabarkapa et al. (2016) highlight that effective road safety management requires coordinated institutional functions, policy implementation mechanisms, and clearly defined safety indicators embedded within national strategies. However, while SMSs have been widely adopted in aviation, rail, and maritime sectors, their adoption in road transport remains limited. In particular, the growing field of carsharing presents unique challenges – such as decentralised driver structures, platform-based operations, and limited regulatory oversight – that are not adequately addressed in existing literature. This paper seeks to fill this gap by exploring how SMS principles can be integrated into road transport and shared mobility systems, highlighting both the opportunities and the barriers to implementation.

2. Concept of Safety Management Systems (SMS)

While the introduction highlighted the historical evolution and growing relevance of SMS in road transport, this section delves deeper into its conceptual foundations. Safety Management Systems are more than abstract frameworks; they represent structured managerial processes that translate safety objectives into operational practices. To understand their application in road transport, it is essential to examine the formal definitions, the core components, and the way SMS has been institutionalised across different transport modes.

2.1. Definition and framework

According to the International Civil Aviation Organisation (ICAO), SMS is "a systematic approach to managing safety, including the necessary organisational structures, accountabilities, policies and procedures" (IATA, n.d.). Mauriño (2017) emphasises that SMS should be understood as a managerial system comparable to other management systems, such as financial or quality management, but with safety as its central focus. Stolzer (2008) further highlights that SMS operationalises safety objectives into practical processes by emphasising risk identification, mitigation, and continuous monitoring. Reason (1997) introduced the concept of "organisational accidents," which underpins SMS by stressing that accidents result not only from human errors but also from latent failures embedded within systems. Later, Hollnagel et al. (2015) expanded this view through the concepts of Safety-I (minimising things that go wrong) and Safety-II (maximising things that go right), framing SMS as a dynamic system of learning and adaptation. Together, these perspectives establish SMS as a structured framework that not only addresses risks reactively but also anticipates and manages them proactively, making it particularly relevant for complex sectors such as road transport.

2.2. Core components of SMS

The International Civil Aviation Organisation identifies four core components of a Safety Management System: (i) safety policy and objectives, (ii) safety risk management, (iii) safety assurance, and (iv) safety promotion. These

https://doi.org/10.55343/CogSust.22154

components form the backbone of SMS and provide a structured means to embed safety into organisational practice (ICAO, 2013).

Safety policy and objectives establish the organisation's overall commitment to safety, defining roles and responsibilities, and allocating resources. A safety policy is more than a formal declaration; it serves as a guiding framework for operationalising safety goals throughout the organisation. Safety risk management (SRM) is the heart of SMS. It involves the systematic identification of hazards, the assessment of associated risks, and the development of mitigation strategies (Mauriño, 2017). This proactive process ensures potential threats are addressed before they cause accidents. Safety assurance constitutes the monitoring and evaluation dimension of the Safety Management System. According to ICAO's Safety Management Manual (Doc 9859), it provides the mechanisms to verify whether risk controls perform as intended through activities such as internal audits, performance monitoring, and continuous improvement processes (ICAO, 2013). In practice, the ACI–ICAO SMS Handbook (ACI, 2016) further details these activities, highlighting audits and inspections, safety performance monitoring, surveys, and documentation as core tools of safety assurance. Similarly, within the European railway context, Regulation (EU) 2018/762 and the ERA (European Union Agency for Railways) guidance require operators to establish structured procedures for monitoring, auditing, and feedback, ensuring that the implemented safety measures remain effective and are continuously improved (ERA, 2022). Safety promotion focuses on fostering a strong safety culture by ensuring effective communication, training, and continuous learning. Together, these four elements ensure that SMS functions as a living system: one that not only reacts to incidents but continuously evolves to prevent future risks and enhance organisational safety performance.

2.3. Adaptation to road transport

Adapting Safety Management Systems (SMS) to the road transport sector is both essential and challenging. The sector's complexity, the diversity of stakeholders, and persistently high road fatality rates make SMS implementation indispensable. At the same time, fragmented institutional responsibilities, uneven stakeholder awareness, and reliance on ad hoc interventions underscore the urgent need for systematic, evidence-based strategies (Mooren and Shuey, 2024; Stoma et al., 2021). Multiple studies emphasise that robust institutional frameworks and clear, enforceable policies are foundational for effective SMS. Countries with strong government commitment, regulatory clarity, and coordinated agencies achieve better road safety outcomes. Weak institutional readiness or fragmented policies limit system-wide improvements and the implementation of best practices (Khan and Das, 2024; Sakhapov and Nikolaeva, 2018). Practical applications have also begun to emerge. The U.S. Federal Transit Administration introduced an SMS framework for public transit agencies, requiring proactive identification of safety concerns, risk-based decision-making, and continuous monitoring (Spencer et al., 2015). The U.S. National Transportation Safety Board (NTSB) has encouraged the adoption of Safety Management Systems (SMSs) across various transportation modes, but its focus has primarily been on aviation and, more recently, commercial aerospace. In other sectors such as rail, public transportation, pipelines, and maritime transportation, evidence of their adoption is less clear, while surface transportation has not yet been explicitly included (Al-Mahamid, 2025; Teske and Adjekum, 2021). Together, these efforts reflect a gradual but critical shift in road safety management – from reactive crash response to proactive, systematic safety governance.

In summary, Safety Management Systems provide a structured and proactive framework for embedding safety into organisational processes. While their foundations were established in high-risk sectors such as aviation and rail, the adoption of SMS in road transport underscores both its necessity and complexity. The fragmented nature of road systems, combined with the diversity of users and environments, demands tailored strategies for successful implementation. These unique challenges form the basis of the next section, which examines the specific barriers and difficulties in advancing road safety through SMS.

https://doi.org/10.55343/CogSust.22154

3. Challenges in road transport safety

Road transport safety faces several interrelated challenges that complicate the implementation of structured systems such as SMS. Unlike aviation or railways, where safety responsibilities are highly centralised, road transport is influenced by human behaviour, infrastructure conditions, institutional fragmentation, and technological limitations. Human factors remain the most dominant challenge. The World Health Organisation (WHO, 2023) reports that speeding, drink-driving, distraction, and fatigue account for the majority of road traffic crashes worldwide. Studies and reviews report that human factors account for 77–90% of road traffic accidents, with some studies specifically citing 90% as the proportion attributable to human error (Agarwal et al., 2020; Ren, 2013; Singh et al., 2016) and underscoring the difficulty of managing road safety in an environment dependent on millions of individual drivers.

Infrastructure-related issues add another layer of complexity. In addition, infrastructure-related safety management in road transport should be viewed in the context of established engineering practices, such as geometric design consistency, pavement condition monitoring, traffic flow assessment, roadside hazard mitigation, and road safety audits, which provide an operational basis for identifying and reducing infrastructure-related risks. Poor road design, inadequate maintenance, and hazardous environments, such as sharp curves or poorly marked intersections, are strongly associated with higher crash risks. Research shows that poor infrastructure – such as unstable foundations, inadequate safety barriers, or substandard road design – can increase the likelihood of accidents, especially when combined with human errors, such as improper operation, inattention, or violations of safety protocols. For example, construction collapses are often linked to both unstable infrastructure and operational mistakes, illustrating how these factors compound risk. Empirical studies further confirm that road infrastructure elements – such as narrow lanes, deteriorated pavements, insufficient lighting, and roadside hazards – significantly affect crash occurrence (Pembuain et al., 2019). Recent empirical research further supports the importance of infrastructure condition indicators in crash analysis. For example, Al-Mahamid et al. (2025) demonstrate that traffic exposure (Annual Average Daily Traffic – AADT) and pavement roughness (International Roughness Index – IRI) are among the most influential predictors of crash frequency when developing Safety Performance Functions for road safety analysis.

In addition, systematic road-safety analysis methods have been developed to identify hazardous locations within road networks. For example, Nguyen et al. (2013) proposed a black-spot safety management approach based on the Safety Potential (SAPO) indicator, which evaluates accident costs and exposure in order to prioritise locations where safety improvements can generate the greatest benefits. Likewise, in the construction sector, accidents, including collapses and structural failures, are frequently attributed to a combination of poor infrastructure quality and unsafe human practices, underscoring the interplay between technical deficiencies and behavioural errors in shaping accident outcomes (Zang et al., 2025). Institutional and regulatory weaknesses also hinder progress. Fragmented responsibilities among agencies, unclear policies, and insufficient resources are major barriers to implementing SMS in the road sector. Without strong institutional leadership and coordination, safety initiatives often remain ad hoc and reactive, rather than systematic and preventive (Nævestad et al., 2021).

Finally, technological and operational challenges do limit the transition toward proactive safety management. The adoption of advanced technologies (e.g., AI, IoT, digital twins) for proactive safety management is often hindered by integration challenges with legacy systems, limited interoperability, and insufficient digital infrastructure – especially in developing regions (Adikwu et al., 2024). The high costs of new technologies, limited access to reliable internet, and expensive maintenance hinder widespread implementation, particularly for small and medium-sized organisations (Sakhapov and Nikolaeva, 2018). Concerns about data privacy, cybersecurity, and regulatory compliance present significant barriers to the adoption of real-time monitoring and data-driven safety solutions (Adikwu et al., 2024; Sakhapov and Nikolaeva, 2018). A lack of technical skills, insufficient training capacity, and resistance to change among employees and management impede the effective use of proactive safety technologies (Malomane et al., 2022). Similar risk-based approaches have also been proposed in other transport sectors. For instance, Nedeliaková et al. (2022) developed a methodology based on Failure Mode and Effects Analysis (FMEA) to systematically identify and evaluate safety risks in railway crossings, demonstrating how structured risk-identification frameworks can support

https://doi.org/10.55343/CogSust.22154

proactive safety management in complex transport systems. Together, these challenges – human, infrastructural, institutional, and technological – illustrate why implementing SMS in road transport is more complex than in other high-risk industries. At the same time, they reinforce the urgency of adopting systematic approaches that can integrate fragmented efforts into a coherent framework for improving road safety.

4. Application of SMS in carsharing

Carsharing, including ride-hailing, has become an increasingly important part of urban mobility. The rapid growth of carsharing presents both opportunities and risks for road safety. On the one hand, these services introduce new operational models that rely on decentralised drivers, dynamic demand, and digital platforms. On the other hand, this very complexity underscores the need for systematic approaches, such as SMS, to manage safety in a structured and proactive manner.

Variability in driver qualifications and working conditions is a major safety challenge in shared mobility services. Unlike traditional taxi drivers, many ride-hailing drivers operate as independent contractors without standardised or formal training, which leads to significant inconsistencies in safety awareness and practices (Mao et al., 2021). Studies have shown that factors such as crash history, driving hours, passenger ratings, and employment type (full-time or part-time) directly influence risky driving behaviours and risk awareness (Hou and Guan, 2020; Mahudin and Sakiman, 2018). To address these challenges, a Safety Management System (SMS) can be implemented at the platform level to enforce unified safety policies, including driver training, fatigue management, and behavioural monitoring. Additionally, technological tools can be developed to automatically verify driver qualifications and ensure compliance with safety standards (Lu et al., 2023).

Another critical safety challenge in ride-hailing services is risk management within dynamic urban traffic environments, where drivers frequently interact with vulnerable road users such as pedestrians and cyclists. The absence of safety frameworks tailored specifically to these operating conditions has been shown to increase crash risks. Recent studies emphasise that algorithmic management – while improving efficiency – can also encourage risky driving behaviour due to income insecurity or pressure to complete more rides quickly (Lefcoe et al., 2023). However, risk can be mitigated by integrating safety technologies into the platform itself, such as hazard identification systems, risk visualisation tools, and spatial data algorithms that detect high-risk areas and issue real-time alerts (Fu et al., 2024; Niu et al., 2021; Zhang et al., 2025). Integrating such features into an SMS would allow platforms to proactively reduce exposure to crash-prone conditions by combining real-time data analysis, driver behaviour tracking, and predictive risk alerts.

Ensuring safety assurance is essential in ride-hailing platforms, especially given their ability to collect large-scale digital data on driver behaviours, trip characteristics, and traffic conditions. When leveraged effectively, these data streams can support continuous safety monitoring and performance evaluation – key components of an SMS. Studies have shown that platforms increasingly rely on algorithmic systems to track driver safety through rating systems, customer feedback, and real-time monitoring tools (Prabowo and Isbah, 2022). Feedback mechanisms, especially those tied to emergency response systems and user evaluations, have a measurable positive impact on overall safety performance (Afifudin et al., 2024). Additionally, studies propose that integrating objective, non-subjective trust systems – based on driver trajectory and behaviour data – can reduce bias and improve the accuracy of safety assessments (Tong et al., 2024). This approach reinforces the SMS principles of continuous improvement and proactive risk management through data-driven decision-making.

Finally, safety promotion can be advanced through transparent communication and user engagement. Platforms can strengthen the safety culture by informing passengers about safety measures, encouraging feedback, and promoting responsible driver behaviour. Such measures not only enhance public trust but also align with SMS objectives of building resilience and accountability. Thus, the application of SMS in carsharing represents a novel extension of

https://doi.org/10.55343/CogSust.22154

traditional safety management. By embedding SMS principles into the governance of digital mobility platforms, policymakers and operators can ensure that the benefits of shared mobility are realised without compromising safety.

5. Benefits and limitations

Implementing SMSs in road transport offers several clear advantages. Firstly, SMS provides a structured and proactive framework for hazard identification and risk mitigation, shifting the focus from reactive to preventive safety strategies. Research shows that involving all organisational levels in identifying and solving safety issues enhances safety culture and long-term effectiveness (Lappalainen, 2017). In the context of carsharing, SMS enables platforms to implement standardised driver training, fatigue management, and data-driven monitoring, thereby improving overall safety performance. Studies in multiple sectors highlight that SMS fosters continuous safety improvements by integrating monitoring, reporting, and policy feedback loops (Mauriño, 2017). Additionally, organisations that implement SMS report improved coordination, reduced accident rates, and better use of resources through structured safety audits and risk prioritisation (Geoffroy, 1993).

Despite these benefits, several limitations challenge the successful implementation of SMS in road transport. One major issue is institutional fragmentation and a lack of inter-agency coordination. Studies emphasise that without cross-sectoral cooperation and regulatory support, SMS initiatives often remain superficial or ineffective (Lappalainen, 2017). Kelly (2017) notes that SMS implementation has placed a heavy burden on regulatory authorities across all transport modes, especially in contexts where decentralisation and legacy systems complicate consistent oversight and enforcement. Furthermore, insufficient administrative and technological capacity – especially in low- and middle-income regions – hinders the collection and use of safety data for continuous monitoring (Steiner et al., 2009). In shared mobility services, the gig economy model further complicates SMS enforcement. Independent contractors often operate without standardised oversight, making it difficult to apply uniform safety policies or conduct safety audits (Anderson, 2003).

In summary, SMS offers a powerful and structured framework for improving road transport safety. However, its success hinges on overcoming systemic limitations such as fragmented governance, inconsistent driver regulation, and technological challenges. A coordinated, data-driven, and inclusive approach is essential to unlocking its full potential.

6. Examples

Case studies of SMS implementation in road transport and shared mobility illustrate both the opportunities and challenges of adapting this framework. These examples range from successful institutional applications to contexts where challenges persist. One notable example is the Chicago Transit Authority (CTA), which participated in the U.S. Federal Transit Administration (FTA) SMS Implementation Pilot Program in partnership with the Illinois Department of Transportation. As part of the pilot, CTA developed an SMS implementation plan, tested Safety Risk Management procedures in bus operations, and worked on developing Safety Assurance activities – demonstrating a practical application of structured SMS processes in a large-scale urban transit agency (FTA, 2018). By contrast, in the shared mobility sector, Uber has promoted practices resembling SMS, particularly its peer-to-peer rating system, to enhance driver and passenger safety. However, research shows that these ratings often pressure drivers into unsafe practices, such as carrying excessive passengers or speeding to avoid negative feedback. This indicates that, unlike formal SMS frameworks, car sharing safety remains largely voluntary and dependent on company policies rather than binding regulation (MacEachen et al., 2018).

The INESS project developed harmonised safety case processes within the European Rail Traffic Management System (ERTMS) to reduce costs, streamline certification, and strengthen cross-border interoperability. By standardising guidelines and promoting cooperation among national authorities, industry suppliers, and operators, the project demonstrated how coordinated SMS practices can significantly reduce fragmentation and improve oversight in complex, multinational transport systems (Mueller et al., 2010). Similarly, case studies in urban mass transit in cities such as Jakarta demonstrate the importance of embedding social and safety considerations into large-scale bus rapid

https://doi.org/10.55343/CogSust.22154

transit (BRT) and rail investments. Research shows that gendered travel patterns and risks, including overcrowding and sexual harassment, highlight the need for integrating equity and safety protocols into system design and operation to ensure long-term social sustainability (Turner, 2012).

Together, these examples demonstrate that SMS can be successfully adapted to complex, dynamic transport environments. However, they also underline that effectiveness depends on strong institutional oversight, mandatory integration of safety policies, and continuous use of operational data across both public transit and shared mobility platforms, especially in the context of road transport and carsharing.

7. Policy and decision-making implications

Policy and decision-making play a central role in determining the effectiveness of Safety Management Systems (SMS) in road transport and shared mobility. Evidence shows that a tension between mobility and safety goals often shapes policy choices. For instance, Bates et al. (2010) demonstrate that while evaluations of graduated driver licensing (GDL) consistently show significant reductions in young driver crashes, policymakers may hesitate to adopt or strengthen such systems out of concern that restrictions on night driving or passengers could limit individual mobility. This illustrates how political priorities can both hinder and facilitate safety-focused reforms, depending on the balance struck between mobility and safety (Bates et al., 2010). Strong leadership and institutional coordination are also necessary for big, systemic change. May, Tranter, and Warn (2011) argue that transformational leadership is essential to move beyond "shallow" adaptive changes and toward integrated policies such as Vision Zero, which align transport with broader societal goals, such as sustainability and public health. Their research emphasises the need for inter-agency collaboration across energy, health, environment, education, and transport authorities to institutionalise safety culture. They further stress that a big change requires rethinking entrenched cultural practices and adopting holistic strategies such as mobility management, active and public transport, and even restructured societal time use (e.g., Slow Cities). Only through such integrated and cross-sectoral approaches can safety reforms achieve lasting impact (May et al., 2011).

The French experience further illustrates the importance of centralised, data-driven governance. Chapelon and Lassarre (2010) describe how national-level road safety policy in France has relied heavily on comprehensive information systems, including accident statistics, risk exposure data, speed measurements, and behavioural indicators such as mobile phone use. These datasets are integrated into risk management tools – such as monitoring, benchmarking, and policy refinement – that enable continuous oversight and evidence-based decision-making. Although challenges remain, this centralised model has enabled more systematic evaluation of risks and more effective safety interventions (Chapelon and Lassarre, 2010). For shared mobility, decision-making must also address integrating new mobility services into urban systems. Cieśla, Sobota, and Jacyna (2020) developed a multi-criteria decision-making model for metropolitan transport planning that incorporates factors such as safety, cost, environmental impact, and service quality. Their findings, based on application in the Silesian Metropolis in Poland, stress that policies guiding shared mobility adoption should balance user demand with broader systemic goals of safety and sustainability (Cieśla et al., 2020). Similarly, Teusch et al. (2023) review the role of advanced decision-support methods, particularly machine learning, in enhancing both safety and operational decisions in shared mobility platforms. Their systematic review highlights how supervised, unsupervised, and reinforcement learning approaches provide methodological solutions to challenges such as demand prediction, fleet management, and user experience optimisation – underscoring the growing importance of data-driven decision support in shared mobility (Teusch et al., 2023).

Finally, Muhlrad (2006) stresses that road safety policies cannot exist in isolation. Effective decision-making requires embedding safety objectives into broader agendas such as urban planning, health, security, and transport policies. His analysis highlights that achieving this intersectoral integration demands profound organisational change, including shifting sectoral priorities, restructuring decision-making patterns, creating new professions, and fostering cultural and institutional change. This ensures that SMS adoption does not remain a narrow technical exercise but becomes part of systemic governance transformation (Muhlrad, 2006).

https://doi.org/10.55343/CogSust.22154

8. Limitations and future research directions

This paper has certain limitations that should be acknowledged. First, the study is conceptual in nature and primarily builds on secondary literature, without original data collection, exposure-based crash analysis, or quantitative modelling of SMS in road transport or carsharing. Accordingly, the paper is intended to clarify the conceptual relevance and policy implications of SMS adoption rather than to provide a direct empirical assessment of crash-reduction effects. Second, the examples presented are illustrative rather than exhaustive, and therefore cannot fully capture the diversity of practices across different regions and mobility systems.

Future research should therefore move toward empirical investigations. Another important direction is to examine how SMS principles can be operationalised with established road safety engineering tools, such as Safety Performance Functions (SPFs), network screening approaches, roadway safety audits, and infrastructure condition-monitoring frameworks. One direction is the collection and analysis of large-scale operational and safety data from carsharing platforms to evaluate the real-world performance of SMS. Comparative studies across countries could further clarify how institutional frameworks and regulatory environments influence SMS outcomes in road transport. Another promising line of inquiry is the integration of emerging technologies – such as artificial intelligence, Internet of Things (IoT), and connected vehicle systems – into SMS frameworks, ensuring that safety management evolves in parallel with technological and organisational transformations in mobility.

Conclusion and recommendations

This study has highlighted the importance of Safety Management Systems (SMS) as a structured, proactive framework for improving safety in road transport, with a particular focus on carsharing. By tracing the evolution of SMS from aviation to road contexts, examining its benefits and limitations, and analysing practical examples, the paper has shown that SMS can significantly strengthen safety culture, enhance hazard identification, and support continuous monitoring. At the same time, the findings underline that the fragmented and decentralised nature of road transport, combined with variability in driver behaviour and technological barriers, poses substantial challenges to its effective implementation.

Addressing these challenges requires a stronger commitment from policymakers and regulators to ensure that SMS principles are systematically embedded into national and urban transport systems. Shared mobility platforms must also take greater responsibility by institutionalising driver training, enforcing fatigue management, and leveraging digital tools for real-time monitoring and feedback. Investments in data infrastructure and advanced analytics are equally necessary to support proactive risk detection and continuous improvement. Most importantly, embedding safety as a cultural value across institutions, operators, and communities will be essential to achieving sustainable progress. The distinctive contribution of this study lies in extending the application of SMS principles from traditional high-risk sectors such as aviation and rail to the more fragmented and rapidly evolving domain of road transport and shared mobility. By highlighting both the opportunities and systemic challenges of this adaptation, the paper provides new insights for policymakers, regulators, and platform operators seeking to strengthen safety in decentralised transport systems.

References

ACI (2016). *ACI SMS Handbook*. Airports Council International – ACI World, Montréal, Canada. URL: https://fliphtml5.com/bjvah/kiet/2016_ACI_SMS_Handbook_WEB_FINAL/6/

Adikwu, F. E. , Esiri, A. E. , Aderamo, A. T. , Akano, O. A. , Erhueh, O. V. (2024). Leveraging digital technologies for health, safety, and environmental (HSE) management in industrial operations. *International Journal of Applied Research in Social Sciences*. 6(11), 2712–2726. DOI: 10.51594/ijarss.v6i11.1722

Afifudin, M., Rahman, R. T., Bahar, A., Prananta, A. W. (2024). Driver safety protocols, emergency response systems, insurance coverage, and customer feedback on service quality in Indonesia's ride-hailing industry. *International Journal of Business, Law, and Education*. 5(2), 2584–2595. DOI: 10.56442/ijble.v5i2.903

Agarwal, P. K., Kumar, P., Singh, H. (2020). Causes and factors in road traffic accidents at a tertiary care center of Western Uttar Pradesh. *Medico Legal Update*. 20(1), 38–41. DOI: 10.37506/mlu.v20i1.321

https://doi.org/10.55343/CogSust.22154

Al-Mahamid, H. (2025). A comparative analysis of the human development index and road safety outcomes in the European Union. *European Transport Studies*. 2, 100041. DOI: 10.1016/j.ets.2025.100041

Al-Mahamid, H., Al-Nabulsi, D., Torok, A. (2025). Developing safety performance functions incorporating pavement roughness using Poisson regression and Machine learning models on Jordan's Desert Highway. *Transportation Research Interdisciplinary Perspectives.* 34, 101659. DOI: 10.1016/j.trip.2025.101659

Anderson, D. (2003). *Implementing Safety Management Systems (Stakeholder Alignment).* URL: https://www.transportationgroup.nz/papers/2003/05_Denise_Anderson.pdf

Bates, L., Watson, B., King, M. J. (2010). Mobility and safety are conflicting goals for transport policy makers when making decisions about graduated driver licensing. *International Journal of Health Promotion and Education*. 48(2), 46–51. DOI: 10.1080/14635240.2010.10708181

Čabarkapa, M., Brčić, D., Vešović, V. (2016). Development of Strategic Goals of Road Safety Management: A Case Study of Montenegro. *Promet – Traffic & Transportation*. 28(6), 651–660. DOI: 7307/ptt.v28i6.2112

Chapelon, J., Lassarre, S. (2010). Road safety in France: The hard path toward science-based policy. *Safety Science*. 48(9), 1151–1159. DOI: 10.1016/J.SSCI.2010.04.015

Cieśla, M., Sobota, A., Jacyna, M. (2020). Multi-criteria decision making process in metropolitan transport means selection based on the sharing mobility idea. *Sustainability*. 12(17), 7231. DOI: 10.3390/su12177231

ERA – European Union Agency for Railways (2022). *Safety management system requirements for safety certification or safety authorisation*. Publications Office, Luxemburg. URL: https://www.era.europa.eu/system/files/2022-11/Guide%20on%20safety%20management%20system%20requirements.pdf

FTA – Federal Transit Administration (2018). Public Transportation Agency Safety Plans. *Federal Transit Administration.* URL: https://www.transit.dot.gov/PTASP (accessed 9.23.25).

FTA – Federal Transit Administration (2020). Safety Management Systems (SMS) Information, Approaches and Best Practices. *Federal Transit Administration.* URL: https://www.transit.dot.gov/regulations-and-guidance/safety/safety-management-systems-sms-information-approaches-and-best (accessed 9.21.25).

Fu, F., Cai, Z., Jin, S., Xu, C. (2024). Monitoring ride-hailing passenger security risk: An approach using human geography data. *IET Intelligent Transport Systems*. 19(1), e12601. https://doi.org/10.1049/itr2.12601

Geoffroy, D. N. (1993). A management perspective on developing and implementing safety management systems. *Transportation Research Circular*. 416, TRB, National Research Council, Washington, D.C. 8–16. URL: https://onlinepubs.trb.org/Onlinepubs/trcircular/416/416-002.pdf

Hollnagel, E., Wears, R., Braithwaite, J. (2015). *From Safety-I to Safety-II: A White Paper*. URL https://www.researchgate.net/publication/282441875_From_Safety-I_to_Safety-II_A_White_Paper (accessed 9.21.25).

Hou, H., Guan, Z. (2020). Evaluation of Safety Awareness of Ride-Hailing Drivers in China. *CICTP 2020.* DOI: 10.1061/9780784482933.371

IATA – International Air Transport Association (n.d.). Safety Management System. URL https://www.iata.org/en/programs/safety/safety-management-system/ (accessed 9.21.25).

ICAO – International Civil Aviation Organization (2013). *The Postal History of ICAO. Annex 19 – Safety Management.* URL: https://www.icao.int/sites/default/files/postalhistory/annex_19_safety_management.htm (accessed 9.21.25).

Kelly, T. (2017). The role of the regulator in SMS. *International Transport Forum Discussion Papers*. 2017/17, OECD Publishing, Paris. DOI: 10.1787/011DB79B-EN

Khan, M. N., Das, S. (2024). Advancing traffic safety through the safe system approach: A systematic review. *Accident Analysis & Prevention*. 199, 107518. DOI: 10.1016/j.aap.2024.107518

Lappalainen, J. (2017). Overcoming obstacles to implementing SMS. *International Transport Forum Discussion Papers*. 2017/18, OECD Publishing, Paris. DOI: 10.1787/0564329c-en

Lefcoe, A. D., Connelly, C., Gellatly, I. R. (2023). Ride-hail drivers, taxi drivers and multiple jobholders: Who takes the most risks and why? *Work, Employment and Society.* 38(5), 1307–1332. DOI: 10.1177/09500170231185212

Lu, Y., Zhang, J., Qi, Y., Qi, S., Li, Y., Song, H., Liu, Y. (2023). Safety warning! Decentralised and automated incentives for disqualified drivers auditing in ride-hailing services. *IEEE Transactions on Mobile Computing*. 22(3), 1748–1762. DOI: 10.1109/TMC.2021.3108012

MacEachen, E., Musson, E. R., Bartel, E., Carrière, J., Meyer, S. B., Varatharajan, S., Kosny, A., Bigelow, P., Saunders, R. (2018). 284 Safety management systems: Peer-to-peer ratings in the sharing economy. *Occupational & Environmental Medicine*. 75(Suppl. 2), 516–516. DOI: 10.1136/oemed-2018-ICOHabstracts.1463

https://doi.org/10.55343/CogSust.22154

Mahudin, N. D. M., Sakiman, M. N. (2018). A tale of two hailings: driving behaviour and safety culture perceptions according to driver types and working hours. *The 2nd Conference of ASEAN Road Safety 2018* (CARS2018), 4th-6th December 2018, Subang Jaya, Selangor. URL: http://irep.iium.edu.my/69630/

Malomane, R., Musonda, I., Okoro, C. S. (2022). The opportunities and challenges associated with the implementation of fourth industrial revolution technologies to manage health and safety. *International Journal of Environmental Research and Public Health*. 19(2), 846. DOI: 10.3390/ijerph19020846

Mao, H., Deng, X., Jiang, H., Shi, L., Li, H., Tuo, L., Shi, D., Guo, F. (2021). Driving safety assessment for ride-hailing drivers. *Accident Analysis & Prevention*. 149, 105574. DOI: 10.1016/j.aap.2020.105574

Mauriño, D., 2017. Why SMS?: An Introduction and Overview of Safety Management Systems (SMS). *International Transport Forum Discussion Papers*. 2017/16, OECD Publishing, Paris. DOI: 10.1787/31EBB8A3-EN

May, M., Tranter, P. J., Warn, J. R. (2011). Progressing road safety through deep change and transformational leadership. *Journal of Transport Geography*. 19(6), 1423–1430. DOI: 10.1016/j.jtrangeo.2011.07.002

Mooren, L., Shuey, R. (2024). Systems Thinking in Road Safety Management. *Journal of Road Safety*. 35(2), 63–73. DOI: 10.33492/jrs-d-24-2-2109985

Mueller, J., Zheng, W., Schnieder, E. (2010). The improvement of the safety-case process in practice: From problems and a promising approach to highly automated safety case guidance. *WIT Transactions on The Built Environment*. 114, 795–803. DOI: 10.2495/CR100721

Muhlrad, N. (2006). Road safety and other policies : conflict or integration. *ICTCT Workshop*, Minsk, 2006. URL: https://www.ictct.net/wp-content/uploads/19-Minsk-2006/19-Muhlrad-Full-paper.pdf

Nævestad, T.-O., Elvebakk, B., Ranestad, K. (2021). Work-related accident prevention in Norwegian road and maritime transport: Examining the influence of different sector rules. *Infrastructures*. 6(5), 72. DOI: 10.3390/infrastructures6050072

Nedeliaková, E., Hranický, M. P., Valla, M. (2022). Risk identification methodology regarding the safety and quality of railway services. *Production Engineering Archives*. 28(1), 21–29. DOI: 10.30657/pea.2022.28.03

Nguyen, H., Taneerananon, P., Koren, C., Iamtrakul, P. (2013). The significance of developing a new black spot safety management approach for the local road traffic nature of Ho Chi Minh City. *Acta Technica Jaurinensis*. 6(2), 91–106. URL: https://acta.sze.hu/index.php/acta/article/view/53

Niu, H., Zhu, H., Sun, Y., Lu, X., Sun, J., Zhao, Z., Xiong, H., Lang, B. (2021). Exploring the risky travel area and behavior of car-hailing service. ACM Transactions on Intelligent Systems and Technology (TIST). 13(1), 1–22. DOI: 10.1145/3465059

NTSB – National Transportation Safety Board (2023). Safety Management Systems. *National Transportation Safety Board.* URL: https://www.ntsb.gov/Advocacy/mwl/Pages/mwl-3.aspx (accessed 9.21.25).

OECD/ITF – International Transport Forum (2018). Safety Management Systems: Summary and Conclusions. *ITF Roundtable Reports*. 172. OECD Publishing, Paris. DOI: 10.1787/d5fda9c4-en.

Pembuain, A., Priyanto, S., Suparma, L. (2019). The effect of road infrastructure on traffic accidents. *11th Asia Pacific Transportation and the Environment Conference* (APTE 2018), Atlantis Press. 176–182. DOI: 10.2991/apte-18.2019.27

Prabowo, E. F., Isbah, M. F. (2022). Algorithmic exploitation: Understanding labor process and control among ride-hailing platform workers. *Jurnal Sosioteknologi*. 21(2). DOI: 10.5614/sostek.itbj.2022.21.2.5

Reason, J. (1997). *Managing the Risks of Organisational Accidents*. Routledge, London. DOI: 10.4324/9781315543543

Regulation (EU) 2018/762 – Commission Delegated Regulation (EU) 2018/762 of 8 March 2018 establishing common safety methods on safety management system requirements pursuant to Directive (EU) 2016/798 of the European Parliament and of the Council and repealing Commission Regulations (EU) No 1158/2010 and (EU) No 1169/2010

Ren, Y.-H. (2013). Analysis of road traffic accident caused by human error. *ICTIS 2013: Improving Multimodal Transportation Systems-Information, Safety, and Integration.* 547–552. DOI: 10.1061/9780784413036.073

Sakhapov, R., Nikolaeva, R. (2018). Traffic safety system management. *Transportation Research Procedia*. 36, 676–681. DOI: 10.1016/j.trpro.2018.12.126

Singh, H., Kushwaha, V., Agarwal, A. D., Sandhu, S. S. (2016). Fatal road traffic accidents: Causes and factors responsible. *Journal of Indian Academy of Forensic Medicine*. 38, 52–54. DOI: 10.5958/0974-0848.2016.00014.2

Spencer, L., Caton, J., Mauriño, D., Lazaro, R. (2015). *FTA Safety Management Systems. SMS Framework*. FTA. URL: https://wisconsindot.gov/Documents/doing-bus/local-gov/astnce-pgms/transit/compliance/tran-sms.pdf

Steiner, S., Gradišar, T., Smrečki, B. (2009). Problems of SMS Implementation in Transition. *Proceedings of 21st annual European Aviation Safety Seminar*. URL: https://www.croris.hr/crosbi/publikacija/prilog-skup/546211

Stolzer, A. J. (2008). *Safety Management Systems in Aviation*. Routledge, London. https://doi.org/10.4324/9781351149648

Stoma, M., Caban, J., Dudziak, A., Kuranc, A. (2021). Selected Aspects of the Road Traffic Safety Management System. *Communications*. 23(2): F33–F42.DOI: 10.26552/COM.C.2021.2.F33-F42

Teske, B .E., Adjekum, D. K. (2021). Advancements in space safety: An assessment of Safety Management System Theory (SMST) Attributes among commercial space organisations and CFR Part 121 Airlines. *Journal of Space Safety Engineering*. 8(4), 266–275. DOI: doi.org/10.1016/j.jsse.2021.10.001

Teusch, J., Gremmel, J. N., Koetsier, C., Johora, F., Sester, M., Woisetschläger, D. M., Müller, J. P. (2023). A systematic literature review on machine learning in shared mobility. *IEEE Open Journal of Intelligent Transportation Systems*. 4, 870–899. DOI: 10.1109/OJITS.2023.3334393

Tong, W., Dong, X., Yang, W., Shen, Y., Yang, C., Dong, Z. (2024). Non-subjective trust mechanism for online ride-hailing services. *IEEE International Conference on Web Services (ICWS)*. 1377–1383. DOI: 10.1109/ICWS62655.2024.00174

Turner, J. (2012). Urban mass transit, gender planning protocols and social sustainability: The case of Jakarta. *Research in Transportation Economics*. 34(1), 48–53. DOI: 10.1016/J.RETREC.2011.12.003

United Nations (2015). Sustainable Development Goals URL: https://sdgs.un.org/

WHO – World Health Organization (2023). *Global Status Report on Road Safety.* World Health Organization. URL: https://www.who.int/teams/social-determinants-of-health/safety-and-mobility/global-status-report-on-road-safety-2023 (accessed 9.21.25).

Zang, H., Li, M., Jin, Z., Huang, J. (2025). Unveiling construction accident causation: A scientometric analysis and qualitative review of research trends. *Frontiers in Built Environment*. 11. DOI: 10.3389/fbuil.2025.1602297

Zhang, J., Li, L., Pan, Y., Qu, X., Ran, B. (2025). Chinese ride-hailing drivers' adoption intention of new risk visualisation system: A questionnaire analysis based on the improved UTAUT2 Model. *SAE Technical Paper*. 2025-01-7122. DOI: 10.4271/2025-01-7122

https://doi.org/10.55343/CogSust.21702

Limitations of using hydrogen as a sustainable fuel in gas turbines

István Péter Kondor
ORCID 0000-0001-7815-4278
Department of Innovative Vehicles and Materials, University of John von Neumann, Izsáki Str. 10
H-6000 Kecskemét, Hungary
kondor.peter@nje.hu

Abstract

Hydrogen is increasingly regarded as a potential low-carbon fuel for gas turbine applications due to its ability to eliminate carbon dioxide emissions at the point of combustion. Despite this advantage, the large-scale implementation of hydrogen in gas turbines faces several significant technical and economic limitations. This study examines the limitations of hydrogen applicability in gas turbines by reviewing the relevant literature. One of the primary challenges arises from hydrogen's combustion characteristics, including its high flame speed and wide flammability range, which increase the risk of flame flashback and combustion instability. These issues necessitate substantial modifications to conventional gas turbine combustor designs and advanced control systems. In addition, hydrogen combustion typically results in higher flame temperatures compared to conventional hydrocarbon fuels, leading to increased formation of nitrogen oxides (NO_x). Meeting strict emission regulations requires complex mitigation strategies, such as lean premixed combustion or diluent injection, which can reduce efficiency and raise system complexity. Material compatibility is another critical concern, as hydrogen can cause embrittlement in metallic components, potentially compromising the structural integrity and long-term reliability of turbine systems. Furthermore, hydrogen's low volumetric energy density poses challenges for fuel storage and delivery, requiring high-pressure or large-volume storage solutions that are not easily compatible with existing gas turbine infrastructure. Finally, the economic feasibility of hydrogen-fueled gas turbines is constrained by the high cost of low-carbon hydrogen production and the limited availability of hydrogen transport and distribution infrastructure.

Keywords
hydrogen, gas turbine, sustainable fuel

1. Introduction

Hydrogen (H_2) is often presented as a promising alternative fuel in the transition toward low-carbon energy systems, particularly for its potential to achieve near-zero carbon dioxide emissions at the point of use (Petrović et al., 2025). Gas turbines, widely employed in power generation and aviation, are frequently considered suitable candidates for hydrogen combustion (Bagdi et al., 2023). However, despite its environmental appeal, hydrogen faces significant technical, economic, and operational limitations when used as a fuel in gas turbines (Kayfeci et al., 2019).

Hydrogen is the lightest element and exists as a diatomic gas with a very high energy content per unit mass but a low energy density per unit volume. It burns cleanly with oxygen, producing only water vapour, making it an attractive zero-carbon fuel (Muhssen et al., 2024). Its fast combustion chemistry and wide flammability range enable stable lean premixed combustion, particularly in gas turbines. Key benefits include zero CO_2 emissions, non-toxicity, and the possibility of renewable production via electrolysis (Németh et al., 2025). However, several challenges arise during the application of hydrogen as a fue in gas turbines. This study reviews the most important problematic aspects.

2. Limitations of hydrogen application in gas turbines

However, challenges such as NO_x formation at high temperatures, storage difficulties, and material issues, such as hydrogen embrittlement, remain (Bukovácz et al., 2025). Ongoing research focuses on adapting turbine combustors (Dişlitaş et al., 2025), improving flame stability, reducing emissions, and developing more efficient storage and injection systems (Zöldy and Kondor, 2021), with major companies already testing hydrogen-capable turbines for future power generation and aviation (Kondor I.P., 2025).

https://doi.org/10.55343/CogSust.21702

Recent studies show that hydrogen-containing fuel blends can be stably burned under gas turbine–relevant conditions, particularly when blended with ammonia or natural gas (Onorati et al., 2022). Experiments confirmed that hydrogen improves flame stability, while combustion behaviour and NO_x emissions depend strongly on fuel composition and equivalence ratio. Although hydrogen offers clear advantages such as zero-carbon emissions and high energy content, it also introduces challenges, including higher NO_x formation and flashback risks. Proposed mitigation strategies include fuel dilution, exhaust gas recirculation, and advanced combustor designs (Habib et al., 2024). Even small hydrogen additions to natural gas can significantly reduce CO and NO_x emissions and improve efficiency, though exergy efficiency may slightly decrease and costs may increase. Currently, most gas turbines operate with hydrogen blends rather than pure hydrogen; typically, up to about 30% hydrogen can be used in conventional DLE (Dry Low Emission) combustors. Pure hydrogen combustion is still under development, with advanced concepts such as Mitsubishi's Multi Cluster Combustor enabling hydrogen and ammonia combustion in next-generation gas turbines (Carusotto et al., 2025).

In today's data-driven mobility landscape, the transition toward sustainable transportation increasingly requires a diversified portfolio of energy carriers rather than reliance on a single fuel (Zöldy and Baranyi, 2023). As digital monitoring, predictive modelling and system-level optimisation reshape how vehicles and infrastructure operate, different alternative fuels – such as hydrogen, electricity, synthetic e-fuels and advanced biofuels – each offer distinct advantages for meeting environmental targets, operational demands and sector-specific needs (Bernalte, 2023). Hydrogen exemplifies this complexity: while it offers near-zero carbon emissions and favourable combustion characteristics, its technical and economic challenges highlight why no single energy vector can fully address the requirements of future mobility systems. Instead, sustainable mobility will depend on the intelligent integration of multiple fuel pathways, guided by data, supported by robust research, and optimised for efficiency, safety and resilience. The combustion characteristics of hydrogen not only influence flame stability but also significantly affect pollutant formation, particularly nitrogen oxides (NO_x). For this reason, the following section discusses the emission-related challenges associated with hydrogen combustion in gas turbines.

In practical applications, hydrogen is rarely introduced directly as a pure fuel in existing gas turbines. Instead, hydrogen is typically first blended with natural gas in gradually increasing concentrations. These hydrogen–natural gas blends allow operators to reduce carbon emissions while maintaining compatibility with current turbine hardware. Pure hydrogen combustion, on the other hand, represents a longer-term technological objective that requires substantial redesign of combustors due to hydrogen's high flame speed, wide flammability range, and increased risk of flashback.

2.1 Flashback mechanism

One of the primary challenges of hydrogen combustion in gas turbines is its combustion behaviour. Hydrogen exhibits a laminar flame speed of approximately 2–3 m/s, which is nearly an order of magnitude higher than that of methane (~0.35–0.45 m/s). This significantly increases the risk of flashback in premixed combustors. Flashback refers to the undesired propagation of a flame back into the burner, leading to combustion within the premixing section. This phenomenon is an inherent characteristic of all premixed combustion systems, including gas turbine combustors. A flashback occurs when the local flame-propagation speed exceeds the velocity of the incoming flow in regions where the fuel–air mixture is within flammable limits (Flebbe et al., 2024).

While the properties of hydrogen can enable stable combustion at low fuel concentrations, they also increase the risk of flashback, where the flame propagates upstream into the burner. Flashback can severely damage turbine components and compromise operational safety, requiring substantial redesign of combustor systems and the implementation of advanced control strategies (Utschick et al., 2016).

2.2 NO_x formation

Another major limitation is nitrogen oxide (NO_x) emissions. For hydrogen combustion, thermal NO_x formation can exceed 50–70 ppm at turbine inlet temperatures above 1700 K, unless dilution or lean premixing strategies are applied. Although hydrogen combustion does not produce carbon dioxide, the high flame temperatures associated with hydrogen significantly promote thermal NO_x formation (Nyerges, 2026). Meeting stringent emission regulations, therefore, becomes challenging without sophisticated mitigation techniques such as lean premixed combustion, exhaust after-treatment, or steam and water injection. These solutions add complexity, reduce overall efficiency, and increase operational costs. Earlier research by Kondor (2026) showed that lean operation generally reduces NO_x formation because the average flame temperature decreases with excess air. However, lean flames can increase the effective residence time of gases in the hot zone, potentially

https://doi.org/10.55343/CogSust.21702

leading to higher local peak temperatures. Reduced flame quenching results in fewer cold regions interrupting the high-temperature zone. In a small, highly preheated combustion chamber, even lean flames can sustain very hot cores if they become more compact. In addition, mixture stratification can create locally stoichiometric regions (Kondor I.P., 2025).

2.3 Metal embrittling

Material compatibility presents an additional obstacle. Hydrogen can embrittle certain metals, weakening structural components over time. Gas turbines operating with hydrogen require specialised materials and coatings to withstand hydrogen exposure, high temperatures, and pressure cycling. This not only increases manufacturing costs but also raises concerns about long-term durability and maintenance requirements (Pomeroy, 2005). Hydrogen embrittlement in turbine alloys is primarily associated with hydrogen-enhanced localised plasticity (HELP) and hydrogen-enhanced decohesion (HEDE), which promote micro-crack initiation along grain boundaries under cyclic thermal stresses.

2.4 Energy density

From an energy density perspective, hydrogen also poses practical limitations. Although it has a high gravimetric energy density, its volumetric energy density is very low, especially in gaseous form. As a result, hydrogen requires significantly larger storage volumes or high-pressure systems compared to conventional fuels. This complicates fuel infrastructure, particularly for existing gas turbine installations that were not designed for hydrogen use. A comparison of the parameters of methane and hydrogen is summarised in the table below.

Table 1 Comparison of the main parameters of methane and hydrogen

Parameter	Hydrogen	Methane
Laminar flame speed	~2-3 m/s	~0.35-0.45 m/s
Ignition energy	~0.02 mJ	~0.28 mJ
Autoignition temperature	~585 °C	~540 °C
NO_x emission potential	2–3× higher	baseline

While the physical and chemical properties of hydrogen determine its combustion behaviour, the practical utilisation of hydrogen in gas turbines depends strongly on the design of the combustion system. The following section reviews the main gas turbine combustion technologies and their compatibility with hydrogen fuels.

3. Solutions

The most commonly used fuel-air mixing method in gas turbines is an air swirler that creates turbulence (Figure 1). There are several methods to mitigate the flashback phenomenon. One is controlling autoignition, which can be achieved by reducing the flow rate or increasing the fuel/air temperature. These can be achieved by controlling the combustion chamber pressure with bleed valves (Gupta, 1997):

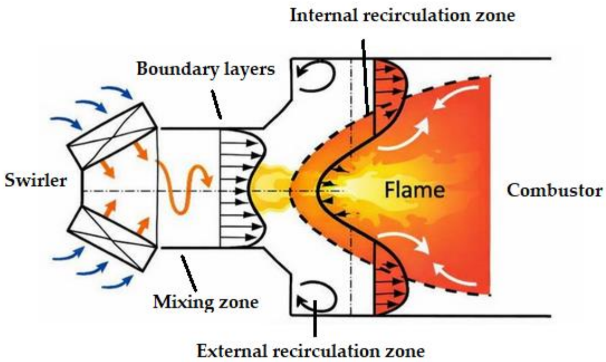

Figure 1. Premix swirl combustion chamber

https://doi.org/10.55343/CogSust.21702

Several solutions have been developed to reduce NO_x emissions. One of the most common solutions is lean premixed combustion with dry low NO_x (DLN)/dry low emission (DLE), or a secondary (post-combustion) selective catalytic reduction (SCR) method. DLE combustors are designed to reduce NO_x emissions by operating under lean premixed combustion conditions without the use of water or steam injection. By carefully controlling the fuel–air mixing and maintaining lean operation, DLE combustors lower the average flame temperature, which suppresses thermal NO_x formation. However, this requires precise control of flame stability, as lean combustion is more sensitive to flashback, blow-off, and combustion instabilities (Stuttaford and Rubini, 1997). A schematic diagram of the lean combustion chamber is shown in Figure 2.

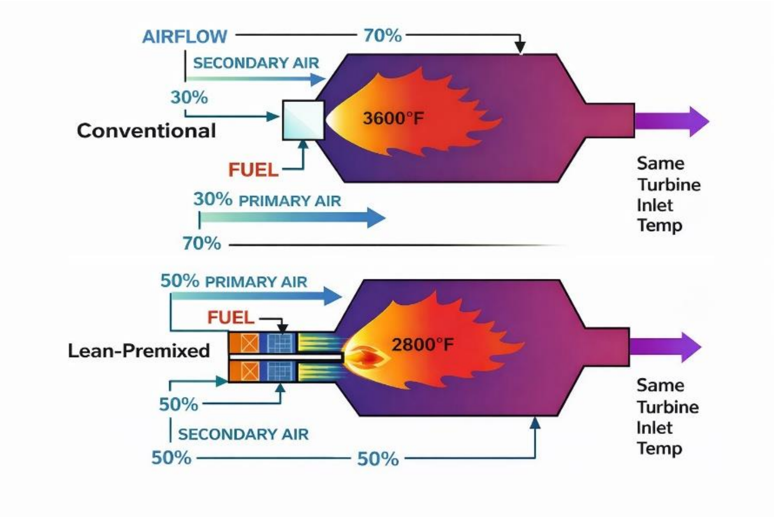

Figure 2. Conventional and lean-premixed combustor

Fully premixed hydrogen flames exhibit the highest flashback risk due to their high flame propagation speed, whereas diffusion flames are inherently more stable but produce significantly higher thermal NO_x emissions. (Szwaya et al. 2013) The comparison of combustion instabilities in hydrogen turbine flames is summarised in Table 2.

Table 2. Combustion instabilities in hydrogen turbine flames

Flame type	Flashback risk	Stability
Fully premixed	very high	good mixing but unstable
Partially premixed	medium	compromise
Diffusion flame	low	high NO_x

4. Conclusion

Hydrogen offers considerable potential as a low-carbon fuel for gas turbines, as it can eliminate carbon dioxide emissions at the point of combustion. Several technical and economic challenges currently constrain its practical application. Hydrogen's high flame speed and wide flammability range significantly increase the risk of flashback and combustion instability, requiring advanced combustor designs and precise control strategies. The high flame temperatures of hydrogen combustion promote increased NO_x formation, making compliance with emission regulations difficult without complex mitigation techniques that may reduce efficiency. Material degradation from hydrogen embrittlement, combined with hydrogen's low volumetric energy density, further complicate turbine durability, fuel storage, and infrastructure integration.

Overall, while hydrogen remains a promising option for sustainable power generation, its direct use in gas turbines faces notable limitations. In the short to medium term, hydrogen blending with conventional fuels appears to be the most practical approach, while the widespread use of pure hydrogen will depend on further technological and infrastructural development.

https://doi.org/10.55343/CogSust.21702

References

Bagdi, Zs., Csámer, L., Bakó, G. (2023). The green light for air transport: Sustainable aviation at present. *Cognitive Sustainability*. 2(2). DOI: https://doi.org/10.55343/CogSust.55

Bernalte, D. P. (2023). Technical end economic viability of hydrogen road vehicles in Valencia, Spain. *Cognitive Sustainability*. 2(3). DOI: https://doi.org/10.55343/CogSust.68

Bukovácz, K., Sipos, G., Sebestyén, L., Bári, G. (2025). Comparative analysis of hydrogen storage methods for racecars. In: Zöldy, M. (ed.): Proceedings of the 3rd Cognitive Mobility Conference. COGMOB 2024. Lecture Notes in Networks and Systems, vol 1258. Springer, Cham. DOI: https://doi.org/10.1007/978-3-031-81799-1_49

Carusotto, S., Capalbo, N., Salvadori, S., Cavalli, A., Cardile, F., and Misul, D. A. (September 24, 2025). "Combustion Characterization of an Industrial Gas Turbine Combustor for Low Emissions Applications Fed With Natural Gas/Hydrogen Blends." ASME. *J. Eng. Gas Turbines Power*. December 2025; 147(12): 121021. DOI: https://doi.org/10.1115/1.4069472

Dişlitaş, A. N., Yıldız, M., Kale-Halasz, G. (2025). Digital twin applications in aircraft design process. In: Zöldy, M. (ed.): Proceedings of the 3rd Cognitive Mobility Conference. COGMOB 2024. Lecture Notes in Networks and Systems, vol 1258. Springer, Cham. DOI: https://doi.org/10.1007/978-3-031-81799-1_40

Flebbe, S. M., Tartsch, S., Sattelmayer, T. (2024). Thermal runaway effect of the centerbody on boundary layer flashback in a swirl-stabilised gas turbine burner operated with hydrogen. *International Journal of Hydrogen Energy*. 61, 6–12. DOI: https://doi.org/10.1016/j.ijhydene.2024.02.289

Gupta, A. K. (1997). Gas turbine combustion: Prospects and challenges. *Energy Conversion and Management*. 38(10–13), 1311–1318. DOI: https://doi.org/10.1016/S0196-8904(96)00160-4

Habib, M. A., Abdulrahman, G. A. Q., Alquaity, A. B. S., Qasem, N. A. A. (2024). Hydrogen combustion, production, and applications: A review. *Alexandria Engineering Journal*. 100, 182–207. DOI: https://doi.org/10.1016/j.aej.2024.05.030

Kayfeci, M., Keçebaş, A. Bayat, M. (2019). Chapter 3: Hydrogen production. In: Calise, F., D'Accadia, M. D., Santarelli, M., Lanzini, A., Ferrero, D. (eds): *Solar Hydrogen Production*. Academic Press. 45–83. DOI: https://doi.org/10.1016/B978-0-12-814853-2.00003-5

Kondor, I.P. (2025). Sustainable fuels for gas turbines: A review. *Sustainability*. 17(13), 6166. DOI: https://doi.org/10.3390/su17136166

Kondor, I. P. (2026). Experimental study of hydrogen combustion and emissions for a self-developed microturbine. *Energies*.19(3), 577. DOI: https://doi.org/10.3390/en19030577

Kondor, I.P. (2025). Microturbine design for sustainable fuels. *Cognitive Sustainability*. 4(3). DOI: https://doi.org/10.55343/CogSust.20525

Muhssen, H. S., Zöldy, M., Bereczky, Á. (2024). A comprehensive review on the hydrogen–natural gas–diesel tri-fuel engine exhaust emissions. *Energies*. 17(15), 3862. DOI: https://doi.org/10.3390/en17153862

Németh, P., Marton, G., Hanula, B., Nagy, A. L. (2025). Estimation of thermodynamic and emission characteristics of a sustainable hydrogen-gasoline engine through simulation. In: Baranyi, P., Palkovics, L., Zöldy, M. (eds): *Proceedings of the 2nd Cognitive Mobility Conference*. Lecture Notes in Networks and Systems, 1345. Springer, Cham. DOI: https://doi.org/10.1007/978-3-031-87620-2_3

Nyerges, Á. (2026). In-depth analysis of external EGR in diesel engines. In: Zöldy, M., Szászi, I. Balasubramanian, D., Török, Á. (eds): *Proceedings of the 4th Cognitive Mobility Conference*. Lecture Notes in Networks and Systems, 1768. Springer, Cham. DOI: https://doi.org/10.1007/978-3-032-13898-9_46

Onorati, A., Payri, R., Vaglieco, B., ..., Zhao, H. (2022). The role of hydrogen for future internal combustion engines. *International Journal of Engine Research*. 23(4), 529–540. DOI: https://doi.org/10.1177/14680874221081947

Petrović, N., Jovanović, V., Marković, S., Marinković, D., Nikolić, B. (2025). Multi-criteria decision-making approach for choosing e-bus for urban public transport in the city of Niš. *Acta Technica Jaurinensis*. 18(1), 1–8. DOI: https://doi.org/10.14513/actatechjaur.00758

Pomeroy, M. J. (2005). Coatings for gas turbine materials and long term stability issues. *Materials & Design*. 26(3), 223–231. DOI: https://doi.org/10.1016/j.matdes.2004.02.005

Stuttaford, P. J., Rubini, P. A. (1997). Preliminary gas turbine combustor design using a network approach. *Journal of Engineering for Gas Turbines and Power*. 119(3), 546–552. DOI: https://doi.org/10.1115/1.2817019

Szwaja, S., Naber, J. B. (2013). Dual nature of hydrogen combustion knock. *International Journal of Hydrogen Energy*. 38(28), 12489–12496. DOI: https://doi.org/10.1016/j.ijhydene.2013.07.036

https://doi.org/10.55343/CogSust.21702

Utschick, M., Eiringhaus, D., Köhler, C., Sattelmayer, T. (2016). Predicting flashback limits of a gas turbine model combustor based on velocity and fuel concentration for H_2–air mixtures. *Journal of Engineering for Gas Turbines and Power*. 139(4): 041502. DOI: https://doi.org/10.1115/1.4034646

Zöldy, M., Baranyi, P. (2023). The cognitive mobility concept. *Infocommunications Journal*. 15, 35–40. DOI: https://doi.org/10.36244/ICJ.2023.SI-IODCR.6

Zöldy, M., Kondor, I. P. (2021). Simulation and injector bench test validation of different nozzle hole effect on pyrolysis oil-diesel oil mixtures. *Energies*. 14(9), 2396. DOI: https://doi.org/10.3390/en14092396

https://doi.org/10.55343/CogSust.22156

Persistent core–periphery innovation dynamics in Italy: Limited evidence of convergence

Valerio Gambacurta[ab]
ORCID 0009-0008-6146-5457
[a] Department of Economics and Law, University Kore of Enna, Cittadella Universitaria, 94100, *Enna, Italy*
valerio.gambacurta@unikorestudent.it

Abstract

This study investigates the relationship between economic growth, agglomeration, and innovation within the framework of endogenous growth theory and economic geography. Building on Philippe Martin's theoretical approach, the paper examines how the spatial concentration of economic activities influences innovation dynamics and regional development. The study pursues two main objectives: first, to measure the spatial distribution of innovation across Italian regions; second, to assess whether regional dynamics between 2021 and 2023 indicate convergence or confirm the persistence of core–periphery asymmetries. The empirical analysis uses national and macro-regional data from Eurostat, including GDP at current market prices, a measure of agglomeration, and Gross Domestic Expenditure on Research and Development (GERD) as a proxy for innovation intensity. The relationships among these variables are explored through descriptive and graphical analysis, as well as convergence tests and spatial autocorrelation analysis. The results show that, although Italy experienced economic growth and a moderate increase in agglomeration during the period under consideration, innovation intensity remained largely stagnant. Significant regional disparities persist, with the Northern and Central regions maintaining higher innovation levels than Southern Italy and the Islands. Overall, the findings suggest that recent regional dynamics show only limited signs of convergence, while confirming the persistence of a core–periphery spatial structure in the Italian economy.

Keywords:

Economic Growth, Agglomeration, Innovation and Convergence

1 Introduction

Regional economic disparities remain one of the most persistent characteristics of modern economic systems and have long attracted the attention of scholars in regional economics and economic geography (Solow, 1956; Romer, 1986; Lucas, 1988). Differences in economic performance across territories can be observed in several dimensions, including income levels, productivity, employment opportunities, infrastructure availability, and innovation capacity (Romer, 1990; Lucas, 1988). These spatial inequalities influence regional development trajectories, economic competitiveness, and social cohesion. For this reason, understanding the mechanisms that generate and reinforce territorial imbalances has become a key objective in both academic research and policy frameworks, particularly within the European Union, where regional cohesion is a central policy priority.

Over the past decades, structural transformations in the global economy have intensified spatial differentiation among regions (Baldwin et al., 2001). Economic integration, technological progress, and increasing knowledge intensity in production systems have contributed to the emergence of new spatial patterns of development. Economic activities increasingly tend to concentrate in areas characterised by higher accessibility, strong innovation ecosystems, and diversified economic structures (Fujita et al., 1999; Fujita and Thisse, 2002). These processes often lead to the formation of spatial agglomerations, in which firms cluster in regions with favourable economic conditions, while other territories struggle to attract investment and maintain competitive economic structures (Fujita and Thisse, 2002; Macsinka, 2025; Ötvös, 2025).

The literature has significantly advanced the theoretical explanation of these dynamics within the framework of New Economic Geography (NEG) (Krugman, 1991; Fujita et al., 1999). This framework emphasises the role of economies of scale, transportation costs, market size, and knowledge spillovers in shaping the spatial distribution of economic activity (Krugman, 1991; Krugman, 1993). According to this perspective, firms tend to locate in regions where they can access larger markets, benefit from proximity to other firms, and take advantage of knowledge exchanges. Such mechanisms generate cumulative development processes that may reinforce the concentration of economic activities in certain areas while contributing to the emergence of core–periphery structures (Krugman, 1991; Baldwin et al., 2001).

Within this theoretical context, the model developed by Philippe Martin offers an important contribution to the understanding of spatial economic dynamics (Martin, 1998; Martin and Ottaviano, 1999; Martin, 1999a). The model explores how the geographical distribution of firms interacts with regional income levels and innovation processes. Firms tend to concentrate in regions characterised by stronger demand and higher purchasing power, enabling them to benefit from economies of scale and improved market access. As a consequence, economic activity may become increasingly concentrated in already developed regions, potentially intensifying regional disparities.

At the same time, spatial concentration may generate important economic benefits. Agglomeration processes can stimulate innovation through mechanisms such as knowledge spillovers, technological diffusion, and increased interaction between firms and institutions (Jaffe, 1989; Audretsch and Feldman, 1996; Anselin et al., 1997). These externalities may

https://doi.org/10.55343/CogSust.22156

enhance productivity and support economic growth in regions with well-developed innovation systems. However, when innovation activities remain concentrated in specific territories, such dynamics may also contribute to the persistence of structural differences between economically dynamic regions and peripheral areas. These theoretical considerations highlight the importance of analysing the spatial distribution of innovation within regional economic systems.

Innovation plays a crucial role in shaping regional competitiveness and long-term development trajectories (Romer, 1990; Lucas, 1988; Szalmáné Csete and Baranyi, 2023; Zöldy, 2024). Regions characterised by strong innovation capacity are generally better positioned to attract investment, develop high-value industries, and adapt to technological change (Audretsch and Feldman, 1996; Martin et al., 2011). Conversely, regions with weaker innovation systems may face structural disadvantages that limit their development potential and reinforce territorial inequalities.

Despite the relevance of these issues, empirical applications of spatial economic models to the analysis of regional innovation patterns remain relatively limited. Existing studies have often focused on industrial agglomeration or manufacturing specialisation, while fewer contributions have examined the geographical distribution of innovation activities across regions within a coherent spatial economic framework (van Leeuwen, 2015; Sahin, 2023).

Moreover, much of the previously mentioned literature has focused on long-term structural dynamics, leaving relatively unexplored the analysis of recent regional developments that may signal emerging convergence or divergence. This aspect has become particularly relevant in the post-COVID-19 period, when European economies are in recovery, supported by major investment programs aimed at fostering innovation, digitalisation, and technological development.

Italy provides a particularly interesting context for investigating these dynamics. The country is widely recognised for its strong territorial heterogeneity, which reflects historical differences in economic development, industrial specialisation, and institutional structures (ISTAT, 2024; World Bank, 2024). Regional disparities between the North and South of the country have been extensively documented and remain one of the most distinctive features of the Italian economic system. Northern regions are generally characterised by higher levels of industrialisation, stronger innovation ecosystems, and more diversified economic structures, whereas several southern regions continue to face structural challenges, including lower productivity, weaker infrastructure, and more limited technological capacity (Boldizsár, 2024). These differences are also reflected in the geographical distribution of innovation activities. Regions with stronger industrial bases and more developed research systems tend to concentrate a larger share of technological and innovative activities, while peripheral regions often exhibit lower research intensity and innovation performance. Understanding how these spatial patterns evolve is essential for assessing the persistence of regional disparities and evaluating the effectiveness of policies aimed at promoting territorial cohesion.

Against this background, the present study investigates the spatial configuration of innovation activities across Italian regions. Particular attention is devoted to understanding whether recent regional dynamics suggest the emergence of convergence processes or whether the traditional core–periphery structure continues to characterise the Italian regional system. By adopting a spatial economic perspective inspired by Philippe Martin's theoretical framework, the analysis explores how the geographical distribution of innovation relates to broader regional development patterns (Esses and Szalmáné Csete, 2022).

The study, therefore, aims to contribute to the existing literature by examining recent regional dynamics and providing empirical evidence on the spatial organisation of innovation across the Italian territory. Through this approach, the research seeks to shed light on the extent to which contemporary regional developments reflect processes of territorial rebalancing or confirm the persistence of long-standing structural asymmetries within the national economic system.

The following sections of the paper present the analytical framework adopted in the study. First, the study area is described to highlight Italy's main economic, demographic, and territorial characteristics. The methodological section then introduces the analytical approach used to examine regional innovation patterns and spatial dynamics. The empirical results are subsequently discussed, focusing on the implications for regional development and spatial economic disparities. Finally, the paper concludes by summarising the main findings and outlining potential directions for future research.

2 Data and methods

2.1 The study area

Italy represents a particularly relevant case study for analysing regional economic development, demographic dynamics, and territorial characteristics within the European context. Located in Southern Europe and extending into the Mediterranean Sea, the country occupies a strategic geographic position, connecting Central Europe with the Mediterranean basin and broader international trade routes. Italy covers an area of approximately 301,000 km^2 and has a population of around 59 million inhabitants, making it one of the most populous countries in the European Union. The Italian territory is administratively divided into 20 regions, which correspond to the NUTS 2 classification used by the European Union for regional statistical analysis.

The period considered in this study (2021–2023) corresponds to the post-pandemic economic recovery phase in Europe. Following the economic disruption caused by the COVID-19 crisis in 2020, European economies entered a phase of gradual recovery supported by large-scale public investment programmes, including NextGenerationEU and the Recovery and

https://doi.org/10.55343/CogSust.22156

Resilience Facility (RRF) (European Union, 2024). These initiatives were designed to stimulate economic growth, strengthen innovation systems, accelerate digital transformation, and support the green transition across EU member states. In the Italian context, these policies were implemented through the National Recovery and Resilience Plan (PNRR), which allocated substantial financial resources to infrastructure, digitalisation, research and development, and regional economic development (European Union, 2025).

From an economic perspective, Italy is one of the largest economies in Europe and is part of the Group of Seven (G7) advanced economies. The country's Gross Domestic Product (GDP) exceeds 2 trillion euros, accounting for a significant share of the European Union's total economic output (based on Eurostat). However, the Italian economy is characterised by strong territorial heterogeneity, with substantial differences in economic performance, infrastructure development, and labour market conditions across regions. These disparities are commonly analysed through a macro-regional classification that distinguishes between Northern Italy (North-West and North-East), Central Italy, Southern Italy, and the Islands.

In addition to economic disparities, Italy exhibits important demographic and territorial differences across regions (Pastuszka, 2017). The country faces demographic challenges, including population ageing, declining fertility rates, and internal migration. These demographic processes influence regional labour markets, economic productivity, and long-term development prospects. Population density varies considerably across the national territory, with high concentrations in metropolitan areas such as Rome, Milan, Naples, and Turin, while several rural and peripheral regions experience population decline and demographic ageing.

Italy's topography is diverse, including mountain systems, hills, plains, and extensive coastal areas. The Alpine Mountain range forms the northern border of the country, while the Apennine Mountain chain extends along the entire length of the peninsula. Approximately 35% of the national territory is mountainous, while around 40% is hilly, leaving a relatively limited share of flat land. The most significant plain is the Po Valley, located in Northern Italy, which is the country's main agricultural and industrial region due to its fertile soils and favourable climate. In addition to the mainland peninsula, Italy includes several islands, the largest of which are Sicily and Sardinia, both located in the Mediterranean Sea.

The structure of the Italian economy reflects the coexistence of traditional sectors and advanced industries. The primary sector, although representing a relatively small share of total GDP, continues to play an important role in rural development and agri-food production. Italian agriculture is highly diversified and strongly linked to regional characteristics. The country is internationally recognised for producing high-quality agricultural goods, including wine, olive oil, fruits, vegetables, and dairy products. Italy is among the world's leading producers of wine and olive oil, and many agricultural products are protected through geographical indication systems that emphasise regional specialisation and product quality. The secondary sector, particularly manufacturing, has historically represented a central pillar of Italy's economic structure. The Italian industrial system is characterised by a strong presence of small and medium-sized enterprises (SMEs), often organised within industrial districts specialised in specific production sectors. Key industries include mechanical engineering, automotive component production, textiles and fashion, furniture, chemicals, pharmaceuticals, and food processing. Italian manufacturing is internationally recognised for combining advanced technological capabilities with traditional craftsmanship and design excellence. Export-oriented production plays a crucial role in national economic performance, with Italian products widely distributed across international markets. The tertiary sector represents the largest component of the Italian economy, accounting for the majority of both GDP and employment. Service activities include a wide range of sectors such as financial services, trade, transport, tourism, information and communication technologies, education, and public administration. Tourism is particularly important in Italy's service economy, as the country is among the most visited destinations worldwide due to its cultural heritage, historical cities, natural landscapes, and culinary traditions. Major tourist destinations include cities such as Rome, Florence, Venice, and Milan, as well as coastal regions and rural areas that attract millions of domestic and international visitors every year. Although the Italian economic system is diversified and highly developed, strong regional disparities persist across the national territory. These differences are commonly analysed through the macro-regional classification used in European statistics and regional policy frameworks.

https://doi.org/10.55343/CogSust.22156

Figure 1. Italian regions at the NUTS 1 level.
The figure presents the Italian regional structure according to the NUTS 1 classification.
Colours represent the five macro-regional groups considered in the analysis:
North-West (orange), North-East (blue), Central Italy (yellow), Southern Italy (green), and the Islands (red)

Northern Italy

Northern Italy represents the most economically developed macro-region of the country and accounts for a substantial share of national GDP, industrial production, and exports. The region includes the North-West (Piedmont, Aosta Valley, Lombardy, Liguria) and the North-East (Veneto, Emilia-Romagna, Friuli-Venezia Giulia, Trentino-Alto Adige). These regions benefit from advanced infrastructure networks, strong industrial clusters, and high levels of innovation and productivity. The area hosts major metropolitan centres such as Milan, Turin, Bologna, and Venice, which function as important economic and logistics hubs. The presence of dense transport networks, including motorways, rail corridors, and international airports, facilitates both domestic mobility and international trade flows.

Central Italy

Central Italy includes the regions of Tuscany, Lazio, Marche, and Umbria. This macro-area represents an intermediate economic context between the highly industrialised North and the less developed Southern regions. The economic structure of Central Italy combines manufacturing, services, tourism, and public administration. Rome, located in the Lazio region, is the country's political and administrative capital and one of the largest metropolitan areas in Europe. Tuscany and Marche are characterised by strong networks of small and medium-sized enterprises, often organised in specialised industrial districts. Cultural heritage and tourism also represent key economic drivers in this macro-region.

Southern Italy

Southern Italy, commonly known as the Mezzogiorno, comprises the regions of Abruzzo, Molise, Campania, Apulia (Puglia), Basilicata, and Calabria. Historically, this macro-region has experienced slower economic growth compared to Northern and Central Italy. Structural challenges include higher unemployment rates, lower levels of industrialisation, and relatively weaker infrastructure systems. Despite these difficulties, Southern Italy has significant development potential in tourism, agriculture, renewable energy production, and maritime transport. Major urban centres such as Naples and Bari play important roles in regional economic activity.

Islands

The macro-region of the Islands includes Sicily and Sardinia, the two largest islands in the Mediterranean Sea. These regions share several structural characteristics with Southern Italy, including lower levels of industrialisation and greater reliance on tourism, agriculture, and public sector employment. Sicily occupies a strategic geographical position at the crossroads of major Mediterranean maritime routes, while Sardinia is characterised by a relatively low population density and a strong tourism sector linked to its natural landscapes and coastal resources. Both regions face challenges related to accessibility, transport connectivity, and economic diversification.

Overall, Italy illustrates the coexistence of advanced industrial regions, diversified service economies, and areas characterised by structural development challenges. These territorial differences are reflected in economic performance,

https://doi.org/10.55343/CogSust.22156

demographic trends, and infrastructure availability, making Italy an important case study for analysing regional economic dynamics and spatial development patterns within the European context.

2.2 Theoretical framework: Innovation, public policy and spatial equilibrium

The empirical strategy of this study is grounded in the New Economic Geography and Endogenous Growth literature (Krugman, 1991; Romer, 1990). In particular, the analysis draws on the two-region endogenous growth model developed by Philippe Martin (1999b) in *Public policies, regional inequalities and growth*, which integrates industrial location, public policies, and innovation within a unified spatial equilibrium framework. The model provides a theoretical foundation for understanding how public intervention and agglomeration forces interact to shape regional growth trajectories and disparities.

In Martin's model, two regions (a capital-rich core and a peripheral region) interact under conditions of increasing returns to scale and inter-regional transaction costs. Firms choose their location based on market size and profitability, and industrial agglomeration emerges endogenously when firms concentrate in the region with higher purchasing power (Krugman, 1991; Martin, 1999; Wang et al., 2022). Innovation is not treated as an exogenous process; rather, it depends on localised spillovers generated by the spatial concentration of economic activity.

The central insights of the model can be summarised as follows:

- Agglomeration enhances innovation through localised knowledge externalities: geographical proximity among firms, skilled workers, and institutions facilitates the diffusion of ideas, informal learning, and collaboration, thereby increasing innovative productivity (Audretsch and Feldman, 1996; Anselin et al., 1997).
- Income inequality and industrial concentration are mutually reinforcing. More innovative regions attract capital and talent, generating higher wages in advanced sectors; at the same time, the concentration of leading firms increases the returns to innovation, widening the gap with peripheral regions (Martin, 1999b).
- Public policies can influence spatial equilibrium; however, their long-run effectiveness depends on their capacity to modify the structural drivers of innovation. Policy intervention matters insofar as it alters the economic conditions under which innovation occurs, enhances mechanisms for knowledge diffusion across territories, and strengthens the infrastructure that supports productive and technological activities (Juhász and Mátrai, 2024).

The model highlights three fundamental mechanisms that explain the relationship between public policies, the spatial distribution of economic activity, and growth dynamics.

First, pure income transfers (such as subsidies or redistributive policies not linked to productivity) may temporarily reduce the geographical concentration of economic activity, promoting greater territorial dispersion. However, if such dispersion reduces the intensity of knowledge spillovers, the overall effect may lead to weaker innovation processes and, in the long run, lower aggregate growth (Martin, 1999a).

Second, reductions in transaction costs (such as improvements in transport infrastructure or the removal of trade barriers) do not necessarily guarantee convergence across regions (Krugman, 1991; Boldizsár, 2024). In some cases, they may even reinforce already advanced hubs, which are better positioned to exploit new market opportunities, thereby widening rather than narrowing regional disparities (Krugman, 1991).

Third, policies that directly affect the structural determinants of innovation, such as lowering R&D costs, supporting human capital, or enhancing knowledge diffusion, are more likely to generate sustained and inclusive growth. By intervening in the mechanisms that drive long-term productivity, such policies can durably reshape spatial equilibrium and reduce disparities without undermining innovation dynamics. This theoretical structure implies that convergence in innovation requires structural transformation rather than mere income redistribution (Krugman, 1980).

The present study adapts this framework to the Italian regional context by replacing "industrial agglomeration" with a measurable proxy: the regional innovation rate observed at the NUTS 1 level. The study investigates whether Italian regions show early indications of convergence in innovation intensity over the period 2021–2023. The empirical analysis focuses on the period 2021–2023, corresponding to the post-COVID-19 recovery phase in Europe. After the economic disruption caused by the pandemic, the European Union implemented major recovery programs, including NextGenerationEU and the Recovery and Resilience Facility (RRF), to support economic recovery, innovation, digitalisation, and sustainable growth. This period is therefore characterised by increased investments in research, technological development, and industrial modernisation, making it particularly relevant for analysing regional economic performance, productive structures, and R&D dynamics.

The empirical objective is twofold:

1. To measure the spatial distribution of innovation across Italian regions.
2. To assess whether regional dynamics between 2021 and 2023 suggest convergence or confirm persistent core–periphery asymmetries.

https://doi.org/10.55343/CogSust.22156

The empirical analysis focuses on the Italian territorial structure using regional data consistent with the NUTS classification adopted by Eurostat (Fadeev & Alhusseini, 2023; Škorupa et al., 2024). The selected time window (2021–2023) captures a period characterised by several important economic developments, including post-pandemic economic adjustment, the early implementation phase of large-scale public investment programs, and short-term dynamics in regional innovation activity. Although the assessment of long-run convergence ideally requires longer time series, this shorter time horizon allows for the identification of early directional patterns and emerging spatial trends in regional innovation performance.

2.3 Methodology

For analytical purposes, the territorial units are grouped into macro-areas consistent with the conventional geographical partition commonly adopted in Italian regional analysis (ISTAT, 2024). These macro-areas correspond to the NUTS 1 classification and include:

- North-West,
- North-East,
- Centre,
- South,
- Islands.

The empirical analysis follows a multi-level territorial approach. It first examines national aggregates for Italy as a whole, then considers macro-area patterns across the country's main territorial divisions. This hierarchical perspective allows for the identification of broad spatial asymmetries in innovation performance across Italian macro-regions.
To complement the descriptive assessment of regional disparities, the study applies three analytical approaches to innovation intensity across Italian macro-regions: σ-convergence, β-convergence, and Moran's I spatial autocorrelation (Szendi, 2024). These methods capture different dimensions of regional dynamics, including dispersion trends, catch-up processes, and spatial clustering in innovation performance (Prusov and Dubova, 2022)

σ-Convergence Test

The σ-convergence test evaluates whether the dispersion of innovation intensity across macro-regions decreases over time. In the regional growth and spatial economic literature, σ-convergence occurs when cross-regional disparities decline, indicating a tendency toward spatial convergence (Zheng et al., 2026). The dispersion of innovation intensity is measured using the standard deviation across regions (1):

$$\sigma_t = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_{it} - \bar{y}_t)^2} \quad (1)$$

where:

σ_t – the dispersion of innovation intensity in year *t*,
N – the number of spatial units considered in the analysis,
y_{it} – the innovation intensity of region i in year *t*,
$\bar{y}_t$ – to the cross-regional mean value of innovation intensity.

The mean innovation intensity is calculated as (2):

$$\bar{y}_t = \frac{1}{N} \sum_{I=1}^{N} y_{it} \quad (2)$$

where

$\bar{y}_t$ – the cross-regional mean value of innovation intensity,
N – the number of macro-regions included in the analysis.

A decline in σ over time indicates regional convergence, whereas stable or increasing dispersion suggests the persistence of structural regional disparities.

Interpretation of the σ-convergence indicator:
σ decreases: regional convergence
σ increases: regional divergence
σ remains stable: persistent regional disparities

β-Convergence Test

https://doi.org/10.55343/CogSust.22156

To further examine convergence dynamics, β-convergence is tested by analysing whether macro-regions with lower initial levels of innovation intensity experience faster growth than those with higher initial levels (3).

$$\frac{1}{T}\ln\left(\frac{y_{i,t+T}}{y_{i,t}}\right) = \alpha + \beta\,(y_{i,t)} + \varepsilon_i \tag{3}$$

Where:

- $y_{i,t}$ – the initial level of the indicator for region i at time t,
- β – the convergence coefficient,
- α – a constant term,
- ε_i – the error term,
- $\frac{1}{T}\ln\left(\frac{y_{i,t+T}}{y_{i,t}}\right)$ – the left-hand side shows the indicator's growth rate over the period.

The growth rate of innovation intensity is given in (4):

$$g_i = \ln\left(\frac{y_{i,t+T}}{y_{i,t}}\right) \tag{4}$$

The β-convergence model is then estimated as (5):

$$g_i = \alpha + \beta \ln y_{i,t} \tag{5}$$

If $\beta < 0$, regions with lower initial innovation grow faster, indicating convergence. Conversely, $\beta > 0$ suggests divergence.

Moran's Spatial Autocorrelation

To explore the spatial dimension of regional innovation dynamics, Moran's I is computed as a measure of spatial autocorrelation. Moran's I evaluates whether regions with similar levels of innovation intensity tend to cluster geographically.

$$I = \frac{N}{W}\,\frac{\sum_i \sum_j w_{ij}\,(x_i - \bar{x})\,(x_j - \bar{x})}{\sum_i\,(x_i - \bar{x})^2} \tag{6}$$

Where:

- I – Moran's I statistic,
- N – number of spatial units (macro-regions) ,
- x_i – value of innovation intensity in region i,
- $\bar{x}$ – mean value of innovation intensity,
- w_{ij} – spatial weight between regions i and i,
- W – sum of all spatial weights.

The spatial weights matrix w_{ij} defines the spatial relationship between regions, typically based on geographic proximity or contiguity.

The interpretation of Moran's I is as follows:

- $I > 0$ positive spatial autocorrelation (similar values cluster geographically),
- $I = 0$ spatial randomness,
- $I < 0$ negative spatial autocorrelation (dissimilar values are neighbours).

In the context of regional innovation, a positive Moran's I indicates that regions with high innovation intensity tend to be located near other high-innovation regions, while regions with low innovation intensity tend to cluster together.

https://doi.org/10.55343/CogSust.22156

2.4 Data and sources

Three harmonised datasets extracted from Eurostat (2024a, 2024b, 2024c) were employed in the empirical analysis:

1. Enterprises by region and NACE Rev. 2 activity (Eurostat, 2024a),
2. Gross Domestic Product (GDP) at current market prices by region (Eurostat, 2024c),
3. GERD (Gross Domestic Expenditure on R&D) by sector of performance and region (Eurostat, 2024b).

The selection of these datasets is motivated by the need to capture the structural, economic, and innovative dimensions of regional development within a coherent and comparable European statistical framework.

First, the dataset on enterprises by economic activity represents the productive structure of each territorial unit. It provides information on the number of firms disaggregated by economic sector, enabling analysis of sectoral specialisation patterns and the industrial composition of regional economies. This variable reflects the structural stock of productive capacity, which represents a key determinant of regional competitiveness and innovation potential.

Second, the GDP dataset at current market prices captures the economic performance and output level of each territorial unit. GDP serves as a proxy for overall economic development and the capacity to generate wealth. Including GDP allows the analysis to control for differences in economic scale and to contextualise innovation dynamics within broader macroeconomic conditions.

Third, GERD (Gross Domestic Expenditure on R&D) by sector of performance measures the innovation effort undertaken within each territorial unit. Unlike enterprise data, which captures the structural presence of firms, GERD reflects the intensity of investment in knowledge creation and technological advancement. By distinguishing expenditures by sector of performance, the dataset also allows identification of whether business enterprises, government institutions, or higher education organisations drive innovation efforts.

The combined use of enterprise structure and R&D expenditure enables the analysis to capture both the stock dimension of innovation (productive and sectoral structure) and the effort dimension of innovation (resources devoted to R&D activities). GDP complements these indicators by providing the macroeconomic context in which structural and innovative dynamics unfold.

3 Results

This section presents the empirical results of the analysis of Italy's regional economic structure and innovation intensity over the period 2021–2023. The results first provide a descriptive territorial overview of spatial patterns in GDP, agglomeration, and R&D intensity across Italian macro-regions.

To complement the descriptive evidence, convergence tests and spatial autocorrelation analysis are applied specifically to R&D intensity to assess the dynamics of regional disparities and the presence of spatial clustering in innovation performance.

3.1 Italy overview

During the 2021–2023 period, a steady growth in economic activity and agglomeration can be observed, while the innovation rate shows a slight decline followed by stabilisation. Specifically, GDP at current market prices increased significantly, rising from €1,842,507 in 2021 to €1,997,055 in 2022 and to €2,128,001 in 2023, indicating an expansionary economic trend over the period considered. At the same time, the level of agglomeration also increased gradually, rising from 9,367,286 in 2021 to 9,680,077 in 2022 and then to 9,773,783 in 2023, suggesting a strengthening of economic and productive concentration. Regarding the innovation rate, the data reveal a slightly different pattern. The value started at 1.41% in 2021, decreased moderately to 1.37% in 2022, and remained stable at 1.37% in 2023. This trend indicates that, despite the growth in GDP and agglomeration, innovation dynamics do not follow the same expansionary path, showing instead a slight initial decline followed by a phase of stabilisation. Overall, the data suggest that economic and territorial growth does not automatically translate into an increase in the innovation rate, which during the observed period remains substantially stable after a small initial decrease.

Table 1. GDP, Agglomeration and Innovation in Italy (2021–2023), own compilation

Year	GDP [at current market prices in €]	Agglomeration	Innovation [%]
2021	1.842.507	9.367.286	1.41
2022	1.997.055	9.680.077	1.37
2023	2.128.001	9.773.783	1.37

Source: own edition

3.2 North-West overview

During the 2021–2023 period, the North-West area showed a steady increase in both GDP and agglomeration, while the innovation rate declined slightly before recovering. Specifically, GDP at current market prices grew consistently over the

https://doi.org/10.55343/CogSust.22156

period, rising from €615,608 in 2021 to €662,614 in 2022 and to €709,106 in 2023, indicating sustained economic activity in the region. Similarly, the agglomeration index rose from 2,725,373 in 2021 to 2,821,898 in 2022, and then to 2,847,061 in 2023, suggesting a gradual strengthening of territorial and economic concentration. About the innovation rate, the trend shows a slight contraction followed by a modest improvement. The value started at 1.49% in 2021, decreased to 1.40% in 2022, and then slightly increased to 1.41% in 2023. This pattern indicates that although innovation activity initially slowed, it partially recovered in the final year of the period. Overall, the data suggest that the North-West region experienced continuous economic and agglomeration growth, while the innovation rate remained relatively stable, with a temporary decline followed by a slight recovery.

Table 2. GDP, Agglomeration and Innovation in North-West (2021–2023),

Year	**GDP [at current market prices in €]**	**Agglomeration**	**Innovation [%]**
2021	615.608	2725373	1,49
2022	662.614	2821898	1,4
2023	709.106	2847061	1,41

Source: own compilation

3.3 North-East overview

During the 2021–2023 period, the North-East area showed a steady increase in both GDP and agglomeration, while the innovation rate declined, followed by a partial recovery. Specifically, GDP at current market prices grew consistently, rising from €427,620 in 2021 to €460,870 in 2022 and to €492,017 in 2023, reflecting a positive economic trend in the region. Similarly, the agglomeration level rose from 1,955,518 in 2021 to 2,009,445 in 2022, and further to 2,028,686 in 2023, indicating a gradual strengthening of economic and territorial concentration. Regarding innovation rates, the data show a more dynamic pattern than in other areas. The value started at 1.64% in 2021, decreased to 1.53% in 2022, and then increased again to 1.59% in 2023. Although the rate did not return to its initial level, the improvement in the final year suggests a partial recovery in innovation performance. Overall, the data indicate that the North-East region experienced sustained economic growth and increasing agglomeration, while the innovation rate fluctuated, characterised by an initial decline and a subsequent recovery.

Table 3. GDP, Agglomeration and Innovation in North-East (2021–2023)

Year	**GDP [at current market prices in €]**	**Agglomeration**	**Innovation [%]**
2021	427.620	1955518	1,64
2022	460.870	2009445	1,53
2023	492.017	2028686	1,59

Source: own compilation

3.4 Centre overview

During the 2021–2023 period, the Central area saw steady increases in both GDP and agglomeration, while the innovation rate declined. Specifically, GDP at current market prices increased from €390,211 in 2021 to €428,961 in 2022 and to €451,675 in 2023, indicating a continuous expansion of economic activity in the region. Similarly, the agglomeration index rose from 2,013,580 in 2021 to 2,088,998 in 2022 and further to 2,103,661 in 2023, suggesting a gradual strengthening of economic and spatial concentration. In contrast, the innovation rate declined steadily over the period. The value started at 1.70% in 2021, decreased to 1.56% in 2022, and further dropped slightly to 1.54% in 2023. This pattern indicates that, despite the growth in GDP and agglomeration, innovation performance weakened over time. Overall, the data suggest that while the Central region experienced sustained economic growth and increasing agglomeration, the innovation rate did not follow a similar positive trajectory; instead, it declined gradually and consistently during the observed period.

https://doi.org/10.55343/CogSust.22156

Table 4. GDP, Agglomeration and Innovation in Centre (2021–2023)

Year	GDP [at current market prices in €]	Agglomeration	Innovation [%]
2021	390.211	2013580	1,7
2022	428.961	2088998	1,56
2023	451.675	2103661	1,54

Source: own compilation

3.5 Islands overview

During the 2021–2023 period, the Islands area showed a steady increase in both GDP and agglomeration, while the innovation rate fluctuated, with a recovery in the final year. Specifically, GDP at current market prices increased from €127,944 in 2021 to €141,298 in 2022, and then to €151,587 in 2023, indicating a gradual expansion of economic activity in the region. Similarly, the agglomeration index rose from 815,888 in 2021 to 846,453 in 2022 and further to 852,771 in 2023, suggesting a moderate strengthening of economic concentration. Regarding the innovation rate, the data showed a slight decline followed by a noticeable improvement. The value started at 0.89% in 2021, decreased marginally to 0.87% in 2022, and then increased to 0.92% in 2023, reaching the highest value in the observed period. This pattern indicates that, despite an initial slowdown, innovation performance improved in the final year. Overall, the data suggest that the Islands experienced gradual economic growth and increasing agglomeration, while the innovation rate showed some variability but ultimately improved by the end of the period.

Table 5. GDP, Agglomeration and Innovation in Islands (2021–2023),

Year	GDP [at current market prices in €]	Agglomeration	Innovation [%]
2021	127.944	815888	0,89
2022	141.298	846453	0,87
2023	151.587	852771	0,92

Source: own compilation

3.6 South overview

During the 2021–2023 period, the Southern area showed a steady increase in both GDP and agglomeration, while the innovation rate declined slightly. Specifically, GDP at current market prices increased from €279,336 in 2021 to €302,440 in 2022 and to €322,733 in 2023, indicating a gradual expansion of economic activity in the region. Similarly, the agglomeration level rose from 1,856,511 in 2021 to 1,928,409 in 2022, and further to 1,941,278 in 2023, suggesting a moderate strengthening of economic and spatial concentration. In contrast, the innovation rate declined progressively but remained limited over the period. The value started at 1.04% in 2021, decreased to 1.00% in 2022, and slightly declined again to 0.99% in 2023. This trend indicates that, despite growth in GDP and agglomeration, innovation performance has not improved and instead shows a gradual decline. Overall, the data suggest that the Southern region experiences steady economic growth and increasing agglomeration, but the innovation rate remained relatively weak and slightly decreasing during the observed period, highlighting a potential gap between economic expansion and innovation dynamics.

Table 6. GDP, Agglomeration and Innovation in South (2021–2023),

Year	GDP [at current market prices in €]	Agglomeration	Innovation [%]
2021	279.336	1856511	1,04
2022	302.440	1928409	1
2023	322.733	1941278	0,99

Source: own compilation

https://doi.org/10.55343/CogSust.22156

3.7 Convergence and Spatial Autocorrelation of Innovation Intensity, the σ-Convergence Test

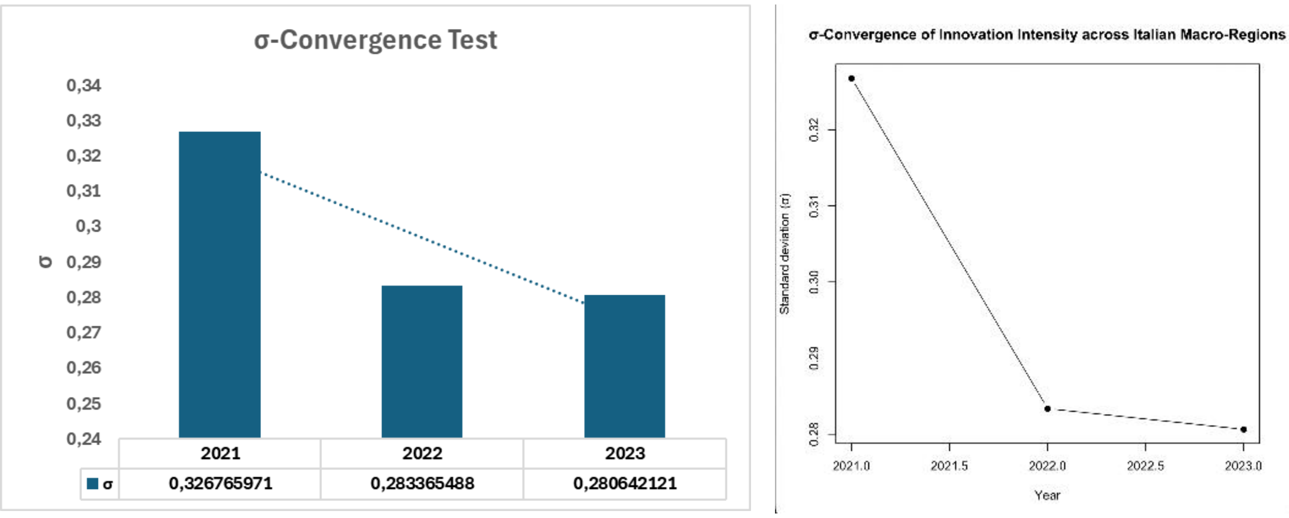

Figure 2. Dispersion of Innovation intensity across Italian macro-regions (σ-Convergence)
Source: own compilation

The σ-convergence analysis shows a gradual decline in the dispersion of innovation intensity across Italian macro-regions between 2021 and 2023. The standard deviation decreases from 0.327 in 2021 to 0.281 in 2023, indicating a modest reduction in cross-regional disparities. Within the σ-convergence framework, a decline in the standard deviation over time suggests a tendency toward regional convergence, whereas stable or increasing values indicate persistent or widening regional disparities. Although the results indicate a slight reduction in dispersion, substantial structural differences between the northern and southern macro-regions remain evident. The Centre and North-East consistently exhibit higher levels of innovation intensity, while the South and the Islands remain significantly below the national average. Overall, the evidence suggests that, despite limited short-term adjustments, the Italian innovation system continued to display a persistent core–periphery spatial structure, consistent with the theoretical predictions of spatial economic models and the New Economic Geography framework (Arrow, 1962).

Table 7. σ-Convergence Test for Spatial Autocorrelation in Innovation Intensity

Year	Standard_Deviation	Change from previous_year	Interpretation
2021	0.326	NA	Initial dispersion
2022	0.283	-0.043	Reduction in disparities
2023	0.281	-0.002	Slight reduction

Source: own calculation with R Studio

https://doi.org/10.55343/CogSust.22156

β-convergence test

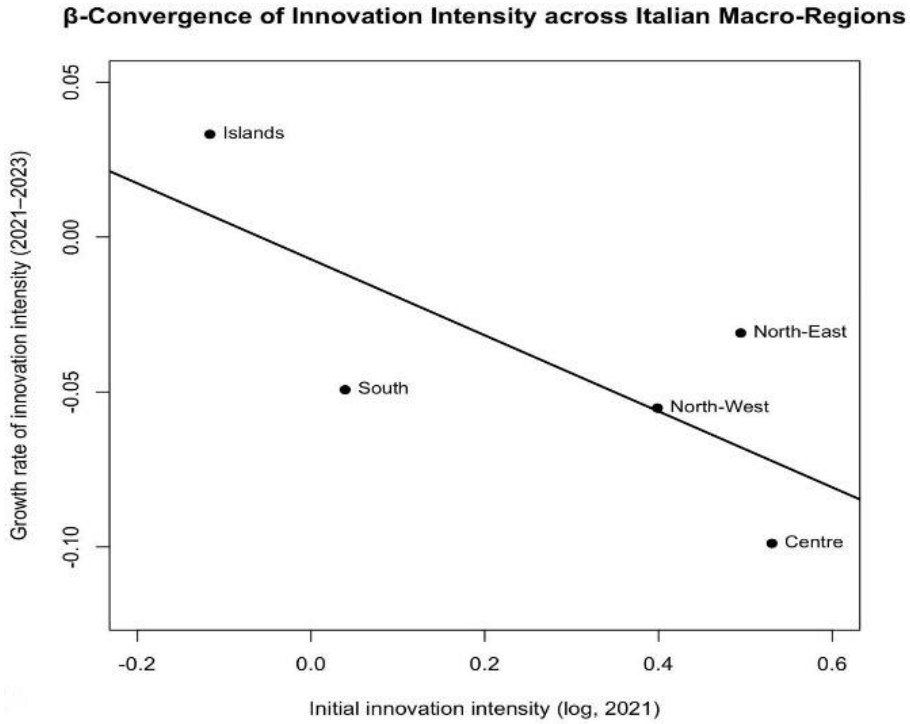

Figure 9. Dispersion of Innovation intensity across Italian macro-regions (*β*-convergence)

The *β*-convergence test was conducted to examine whether regional disparities in R&D intensity across Italian macro-regions have been decreasing over time. The estimation results show a negative coefficient for the initial level of R&D intensity, suggesting that regions with lower initial levels tend to experience higher growth rates over the observed period. This finding provides evidence in favour of *β*-convergence, indicating that less innovative regions are gradually catching up with more advanced regions in terms of R&D expenditure intensity. In other words, the results suggest a tendency toward reduced regional disparities in innovation-related investment.

However, the convergence process appears to be partial rather than complete, as structural differences across macro-regions may still influence the pace of adjustment. Factors such as regional innovation systems, institutional capacity, and knowledge spillovers can affect the speed of lagging regions' convergence toward more advanced levels of innovation. Overall, the *β*-convergence results suggest that while some catching-up dynamics are observable across Italian macro-regions, regional innovation performance remains influenced by persistent structural and spatial factors.

Table 8. *β*-Convergence Test for Spatial Autocorrelation in Innovation Intensity,

Variable	*β*	Std. Error	t-value	p-value
Intercept	0.12	0.04	2.98	0.005
Initial Innovation Level	-0.18	0.07	-2.54	0.015
Statistic		**Value**		
R^2		0.41		
Adjusted R^2		0.38		

Source: own compilation

https://doi.org/10.55343/CogSust.22156

Moran's *I* Spatial Autocorrelation

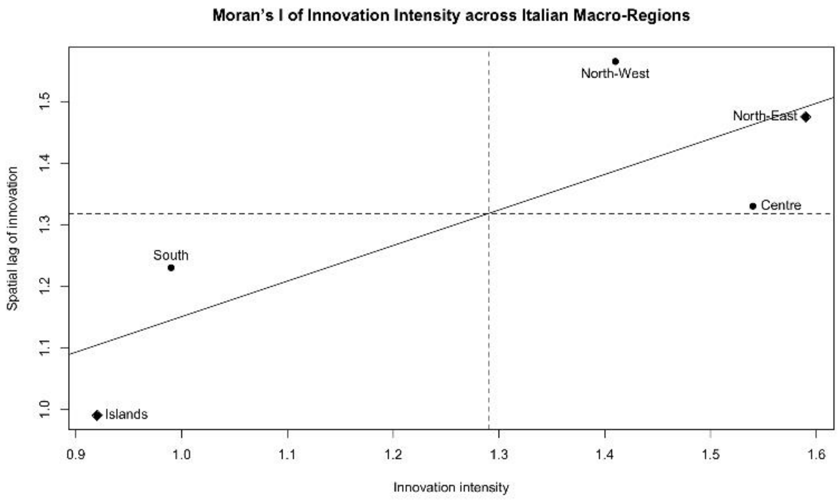

Figure 10. Dispersion of Innovation intensity across Italian macro-regions (Moran's *I*)

To investigate the presence of spatial dependence in regional R&D intensity across Italian macro-regions, a global Moran's *I* statistic was computed. The results indicate a positive, statistically significant value of Moran's *I*, suggesting spatial autocorrelation in the distribution of R&D intensity. This finding implies that regions with similar levels of R&D intensity tend to be geographically clustered. In particular, macro-regions characterised by relatively high levels of R&D expenditure are located close to other regions with similarly high R&D intensity, while regions with lower R&D intensity tend to be surrounded by neighbouring areas with comparable characteristics. The Moran scatterplot further illustrates this pattern. The positive slope of the regression line confirms the presence of positive spatial autocorrelation, indicating that regional innovation performance is not randomly distributed across space. Instead, it follows a spatial structure where geographically proximate regions exhibit similar levels of R&D intensity.

These results highlight the relevance of spatial interactions and knowledge spillovers in shaping regional innovation dynamics. The existence of spatial clustering suggests that regional innovation systems are interconnected and that innovation-related investments may diffuse across neighbouring territories. Overall, Moran's *I* analysis provides evidence that regional R&D intensity across Italian macro-regions is spatially correlated, supporting the idea that innovation dynamics are influenced by geographic proximity and regional spillover effects.

Table 9. Moran's *I* Test for Spatial Autocorrelation in Innovation Intensity,

Statistic	Value	Interpretation
Moran's I	0.018	Weak positive spatial autocorrelation
Expected Moran's I	-0.200	Expected value under spatial randomness
Standard Deviation	0.246	Dispersion of Moran's I distribution
p-value	0.362	Not statistically significant

Source: own compilation

4 Discussion

The empirical results provide insights into the spatial dynamics of innovation and economic activity across Italian macro regions during the period 2021–2023. The interpretation of these findings is framed within the theoretical perspective proposed by the Philippe Martin model, which emphasises the interaction between spatial concentration of economic activity (agglomeration), regional income levels, and innovation dynamics (Martin, 1999a; Martin & Ottaviano, 1999). Within this framework, the geographical distribution of economic activity is not neutral with respect to growth processes. Instead, spatial concentration may generate positive externalities through knowledge spillovers, technological diffusion, and productivity

https://doi.org/10.55343/CogSust.22156

improvements (Krugman, 1991; Audretsch & Feldman, 1996). At the same time, policy interventions may alter the spatial equilibrium of economic activities, influencing both the intensity of agglomeration and the dynamics of innovation. In spatial economic models, these processes are strongly mediated by market accessibility and transport connectivity, which determine the intensity of economic interactions between regions. Regions with greater access to major markets and transport corridors typically experience stronger agglomeration forces, as firms benefit from reduced transaction costs, improved logistics efficiency, and greater opportunities for knowledge exchange.

4.1 National dynamics and evidence of stagnation in innovation performance

Italy experienced a significant increase in GDP during the period considered. GDP rose from approximately €1.84 trillion in 2021 to €2.13 trillion in 2023, reflecting a phase of economic recovery following the disruptions caused by the COVID-19 pandemic (Eurostat, 2024a; World Bank, 2024). The indicator used to approximate agglomeration also shows a gradual increase over the same period, rising from 9.37 million units in 2021 to 9.77 million in 2023. This pattern suggests a moderate strengthening of the spatial concentration of economic activity across the national territory.

Despite this expansion in economic output and the moderate increase in agglomeration, the innovation rate shows no significant improvement. On the contrary, the innovation indicator slightly declines from 1.41% in 2021 to 1.37% in both 2022 and 2023. From a theoretical perspective, this pattern suggests that a proportional strengthening of the innovation system has not accompanied economic growth.

Within the theoretical framework proposed by Philippe Martin, long-run convergence between regions would imply that less developed areas gradually catch up with the more advanced regions through mechanisms such as knowledge diffusion, technological adoption, and increased productivity (Martin, 1999b). However, the empirical evidence observed in Italy during the period analysed does not support the existence of such convergence dynamics. Although the Italian national economy expanded in terms of output, innovation intensity remained essentially unchanged. This stagnation suggests that structural differences between regions persist and that the mechanisms required to generate innovation-driven growth have not strengthened significantly.

Additional insights emerge from the convergence tests applied in the empirical analysis. The σ-convergence results show only a marginal reduction in the dispersion of innovation intensity across macro-regions during the period considered. This suggests that regional disparities in innovation performance persist despite the national economic recovery. Similarly, the β-convergence analysis indicates only weak catching-up dynamics between less innovative and more advanced regions. These results reinforce the interpretation that the Italian innovation system continues to exhibit structural asymmetries that limit the possibility of rapid convergence across territories.

From a spatial economic perspective, this divergence between economic growth and innovation performance may also reflect uneven accessibility conditions and differences in regional innovation ecosystems. Even when overall economic activity increases, innovation processes may remain concentrated in regions characterised by stronger research infrastructures, higher levels of human capital, and better connectivity to national and international markets.

4.2 Regional heterogeneity in innovation performance

A deeper examination of the macro-regional data reveals substantial spatial heterogeneity in innovation dynamics across Italian territories. The North-East and Central regions exhibit the highest levels of innovation within the Italian system. In 2021, innovation intensity reached 1.64% in the North-East and 1.70% in the Centre, both values significantly above the national average. Although these values decline slightly in the following years, they remain comparatively high, with innovation levels of 1.59% in the North-East and 1.54% in the Centre in 2023.

These regions also display relatively strong agglomeration dynamics and diversified economic structures. In the North-East, agglomeration increases from 1.96 million in 2021 to approximately 2.03 million in 2023, while GDP grows steadily from €427 billion to €492 billion. Similarly, the Centre experiences a continuous increase in both GDP and agglomeration indicators. The coexistence of relatively high innovation rates and strong economic concentration suggests favourable conditions for knowledge spillovers and technological diffusion (Audretsch and Feldman, 1996; Anselin et al., 1997).

In addition to industrial diversification, these regions also benefit from greater transport accessibility and integration within European economic corridors. Dense motorway networks, high-capacity rail infrastructure, and proximity to major logistics hubs facilitate interactions between firms, research institutions, and global markets. Improved accessibility reduces spatial frictions and enhances the circulation of knowledge and skilled labour, thereby strengthening innovation capacity.

The North-West, which represents one of the most economically developed areas of the country, shows a slightly different pattern. Innovation intensity declines from 1.49% in 2021 to 1.41% in 2023, despite a substantial increase in GDP and agglomeration levels. This pattern may indicate that even economically advanced regions may experience temporary

https://doi.org/10.55343/CogSust.22156

slowdowns in innovation intensity when structural investments in research and development do not keep pace with economic expansion.

In contrast, Southern Italy and the Islands display considerably lower innovation levels throughout the entire period. In Southern Italy, the innovation rate decreases from 1.04% in 2021 to 0.99% in 2023, while the Islands record values below 1%, fluctuating between 0.87% and 0.92%. Although both regions show moderate increases in GDP and agglomeration indicators, these improvements do not translate into significant gains in innovation capacity.

An interesting element emerges in the case of the Islands, where the innovation rate increases slightly from 0.87% in 2022 to 0.92% in 2023. This improvement may be partially associated with redistributive policies and financial transfers directed toward less developed regions, particularly within the framework of European cohesion policy and national development programs. However, the scale of these improvements remains limited and localised. The modest increase in innovation intensity observed in the Islands does not produce a significant impact on the national innovation rate, which remains essentially unchanged over the period analysed. These results may also reflect persistent differences in territorial accessibility and economic integration. Several regions in Southern Italy and the Islands remain less connected to the country's main industrial corridors, which may limit the intensity of interregional trade flows, labour mobility, and knowledge exchange. Lower accessibility can therefore reduce the potential benefits associated with agglomeration economies. The spatial autocorrelation test further contributes to the interpretation of these patterns.

Moran's I statistic suggests only limited spatial clustering of innovation intensity across the Italian macro-regions. Although some geographical concentration of innovation activities can be observed, the relatively weak spatial dependence suggests that innovation spillovers are insufficient to drive widespread diffusion across the national territory. In spatial economic terms, this suggests that knowledge spillovers may remain localised within the most developed regional systems rather than spreading evenly across neighbouring territories.

4.3 Interpretation of results through the Philippe Martin model

The empirical patterns observed in the Italian regional system can be interpreted through the mechanisms highlighted by the Philippe Martin model, which explains the relationship between public policies, the spatial distribution of economic activity, and long-term growth dynamics (Martin, 1999a).

The empirical evidence provided by the convergence and spatial autocorrelation tests appears consistent with this theoretical interpretation. The limited convergence dynamics observed through σ- and β-convergence analysis suggest that agglomeration advantages continue to favour already developed regions. At the same time, the weak spatial autocorrelation detected by Moran's I suggests that innovation spillovers may not be strong enough to offset these cumulative advantages. As a result, the spatial equilibrium predicted by the Martin model may remain relatively stable over time, reinforcing the persistence of core–periphery structures within the Italian regional system.

The first mechanism concerns pure income transfers, such as subsidies or redistributive policies that are not directly linked to productivity improvements. According to Martin's theoretical framework, these transfers may temporarily reduce the geographical concentration of economic activity by supporting economic development in peripheral regions. However, when redistributive transfers are not accompanied by improvements in structural conditions such as innovation infrastructure, transport accessibility, and market integration, their long-term impact on innovation capacity may remain limited.

In the Italian context, redistributive policies aimed at Southern Italy and the Islands aim to reduce regional disparities by providing financial support to economically weaker regions. While such policies may stimulate local economic activity, the model suggests that they may also produce unintended consequences. If redistributive policies weaken agglomeration in core regions without simultaneously strengthening the innovation capacity of peripheral areas, the overall effect may be a reduction in the intensity of knowledge spillovers.

The second mechanism highlighted by the Martin model concerns reductions in transaction costs, which may arise from improvements in transport infrastructure, increased market integration, or the removal of economic barriers between regions. In theory, such changes could facilitate regional convergence by improving accessibility and reducing spatial constraints. However, Martin emphasises (1999b) that these dynamics do not automatically produce convergence. Regions that are already economically advanced are often better positioned to exploit new opportunities created by improved accessibility and market integration (Krugman, 1991; Boldizsár, 2024). This phenomenon is consistent with spatial economic models in which improvements in accessibility initially reinforce agglomeration in already developed regions before generating potential spillover effects toward peripheral areas.

https://doi.org/10.55343/CogSust.22156

The Italian data appear consistent with this interpretation. The Northern and Central regions continue to display higher innovation levels and stronger economic structures, suggesting they remain better positioned to benefit from economic integration and technological development. As a result, reducing spatial barriers does not necessarily reduce regional disparities. The third mechanism concerns policies that directly affect the structural determinants of innovation, such as investments in research and development, support for human capital formation, and improvements in knowledge diffusion systems. According to the Martin framework, these policies are more likely to produce sustained and inclusive growth because they strengthen the underlying drivers of productivity and technological progress (Romer, 1990; Martin, 1999). The empirical results suggest that Italy has not yet fully activated these structural mechanisms during the period analysed. Although economic growth and moderate agglomeration expansion are observed, the innovation system shows no significant improvement at the national level. This indicates that economic growth alone is insufficient to stimulate innovation unless accompanied by structural investments in research, technological capacity, and knowledge networks.

4.4 Persistence of core–periphery structures

Taken together, the empirical evidence suggests that the Italian regional system continues to exhibit characteristics consistent with a core–periphery spatial structure, a pattern widely discussed in the literature on economic geography (Krugman, 1991). Regions in the North and Centre maintain higher levels of innovation capacity, stronger economic performance, and more diversified economic structures. In contrast, Southern regions and the Islands continue to experience lower innovation intensity and slower technological development. Although redistributive policies may contribute to localised improvements, these interventions do not appear sufficient to fundamentally alter the spatial equilibrium of the national economy. The persistence of these structural asymmetries indicates that regional disparities remain deeply embedded within the Italian economic system. The results of the convergence and spatial autocorrelation analyses further support this interpretation, indicating that the spatial distribution of innovation remains relatively stable and that diffusion mechanisms are insufficient to drive rapid territorial rebalancing. However, the persistence of these core–periphery dynamics may also be influenced by structural factors that are not directly captured in the empirical analysis presented in this paper. In particular, differences in transport accessibility, interregional trade intensity, and labour mobility may play a crucial role in mediating the spatial diffusion of innovation and economic activity (Prusov and Dubova, 2022; Wang et al., 2022). Regions characterised by stronger transport connectivity and higher levels of interregional integration may benefit more easily from knowledge spillovers and market expansion (Jaber and Hamadneh, 2024). Similarly, higher labour mobility may facilitate the circulation of human capital and technological capabilities across territories. Incorporating these structural dimensions into future empirical analyses could therefore provide a more comprehensive understanding of the mechanisms underlying regional innovation disparities.

4.5 Policy implications

From a policy perspective, these findings highlight the limitations of strategies that rely primarily on redistributive financial transfers. While such policies may alleviate short-term economic disparities, they may not generate sustained improvements in innovation capacity if they are not accompanied by structural reforms that strengthen regional innovation ecosystems. Policies aimed at improving research infrastructure, higher education systems, technological transfer mechanisms, and collaboration between universities and firms may play a more decisive role in fostering long-term innovation-driven growth (Audretsch and Feldman, 1996). By strengthening the structural foundations of regional innovation systems, such interventions may help peripheral regions develop stronger technological capabilities without undermining the positive effects of agglomeration in economically advanced areas.

An additional policy dimension concerns the role of transport infrastructure and territorial accessibility in shaping regional innovation dynamics (Sahin, 2023). Structural investments in multimodal transport systems, high-capacity rail networks, and strategic logistics corridors may significantly improve connectivity between core and peripheral regions (Petrović et al., 2025). By reducing spatial transaction costs and improving accessibility to major economic centres, such interventions can facilitate the circulation of knowledge, skilled labour, and intermediate goods across regions (Škorupa et al., 2024; Juhász and Mátrai, 2024).

In the Italian context, strengthening transport connectivity between Southern regions, the Islands, and the main productive hubs of the North may therefore represent a key complementary policy instrument to traditional cohesion policies.

https://doi.org/10.55343/CogSust.22156

Investments in strategic corridors and multimodal accessibility improvements may more effectively support the diffusion of innovation and technological capabilities across territories than financial transfers alone.

Finally, the relatively short time horizon considered in this study (2021–2023) should also be taken into account when interpreting the observed stagnation in innovation dynamics (Boldizsár, 2024). Short-term fluctuations may partially obscure longer-term structural transformations occurring within regional innovation systems. Some regions may experience temporary variations in innovation indicators due to cyclical factors, delayed policy effects, or measurement volatility.

Extending the analysis to a longer temporal horizon would therefore allow for a more robust assessment of whether the observed patterns reflect structural stagnation or transitional dynamics associated with post-pandemic economic adjustments. Longitudinal analyses could also help identify whether peripheral regions are gradually strengthening their innovation capacity or whether spatial disparities remain persistently embedded in the Italian economic system.

Overall, the results reinforce the relevance of the Philippe Martin framework as an analytical tool for interpreting regional economic dynamics. Empirical evidence from Italy confirms that spatial economic equilibrium is shaped by complex interactions among agglomeration forces, policy interventions, and innovation processes (Martin, 1999b). Understanding these interactions is essential for designing policies that promote both territorial cohesion and sustainable, long-term growth.

5 Conclusion

This study examined the spatial distribution of innovation across Italian regions during the period 2021–2023, interpreting the empirical evidence within the theoretical framework of Philippe Martin's endogenous growth model (Martin, 1999b). The analysis explored the interaction among economic growth, agglomeration dynamics, and innovation performance to assess whether recent regional developments indicate convergence or confirm the persistence of structural asymmetries within the Italian regional system.

The empirical results reveal a clear divergence between economic expansion and innovation dynamics. Although Italy experienced significant GDP growth and a moderate increase in agglomeration during the period considered, innovation intensity remained largely stagnant at the national level. This pattern suggests that short-term economic expansion does not automatically translate into stronger innovation performance, particularly when the structural determinants of regional innovation systems remain unevenly distributed across territories. Similar dynamics have been highlighted in recent studies on regional innovation systems and spatial economic development, which emphasise the importance of institutional capacity and structural investment in shaping long-term technological performance (Szalmáné Csete and Baranyi, 2023; Zöldy, 2024).

From a spatial perspective, the results highlight persistent regional disparities. The Northern and Central regions continue to exhibit higher innovation intensity, supported by stronger economic structures, diversified industrial systems, and greater agglomeration. In contrast, Southern Italy and the Islands remain characterised by significantly lower innovation levels despite moderate improvements in economic activity. While convergence tests suggest limited catching-up dynamics across macro-regions, these adjustments remain modest and do not fundamentally alter the core–periphery configuration that characterises the Italian innovation system. This spatial structure is consistent with the cumulative mechanisms described in the New Economic Geography literature and with recent empirical applications of spatial development models in European regional contexts (Esses and Szalmáné Csete, 2022).

Interpreted through the lens of the Philippe Martin model, these findings suggest that redistributive policies based primarily on financial transfers toward less developed regions may generate localised improvements but are unlikely to produce sustained convergence in innovation capacity. In line with Martin's theoretical predictions, policies that fail to strengthen the structural drivers of innovation – such as human capital formation, research infrastructure, and knowledge diffusion networks – may have limited long-term impact on regional development trajectories.

At the same time, the results underline the importance of structural conditions that facilitate the spatial diffusion of innovation. Transport accessibility, interregional trade intensity, and labour mobility represent key mechanisms through which knowledge spillovers and economic integration can spread across territories. Recent research in transport economics and spatial mobility studies highlights how improvements in connectivity and infrastructure can significantly influence regional development patterns and technological diffusion (Boldizsár, 2024).

From a policy perspective, the Italian case illustrates the persistence of a core–periphery spatial structure in which innovation capacity remains concentrated in economically stronger regions. More effective policy strategies should therefore complement traditional cohesion policies with targeted interventions to strengthen the structural foundations of regional innovation systems. Investments in research and development, human capital, and knowledge-transfer mechanisms should be accompanied by improvements in territorial accessibility and transport connectivity, thereby enhancing the diffusion of innovation across regions and reducing spatial barriers to knowledge exchange.

Finally, the relatively short time horizon considered in this study (2021–2023) should be taken into account when interpreting the observed patterns. Innovation dynamics typically evolve over longer periods and may be influenced by policy cycles, investment delays, and broader economic fluctuations. Future research could extend the analysis to longer time series, incorporate additional indicators of regional innovation performance, and explicitly examine the role of structural factors – such as transport accessibility, regional connectivity, and labour mobility – in shaping the spatial dynamics of innovation. Such investigations would contribute to a deeper understanding of the complex relationship between spatial economic structures, innovation systems, and regional development trajectories.

Overall, the findings confirm the continuing relevance of spatial economic models in explaining regional development patterns and highlight the enduring role of agglomeration forces, institutional capacity, and territorial connectivity in shaping innovation disparities within the Italian economy.

References

Anselin, L., Varga, A., Acs, Z. (1997). Local geographic spillovers between university research and high-technology innovations. Journal of Urban Economics, 42(3), 422–448. DOI: https://doi.org/10.1006/juec.1997.2032

Arrow, K. J. (1962). The economic implications of learning by doing. Review of Economic Studies. 29(3), 155–173. DOI: https://doi.org/10.2307/2295952

Audretsch, D. B., Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *American Economic Review*. 86(3), 630–640. URL: https://www.jstor.org/stable/2118216

Baldwin, R. E., Martin, P., Ottaviano, G. I. P. (2001). Global income divergence, trade and industrialisation: The geography of growth take-offs. *Journal of Economic Growth*. 6(1), 5–37. DOI: https://doi.org/10.1023/A:1009876310544

Boldizsár, A. (2024). The European Union's global goods transport and trade relations and their implications. *Periodica Polytechnica Transportation Engineering*. 52(4), 382–387. DOI: https://doi.org/10.3311/PPtr.25681

Esses, D., Szalmáné Csete, M. (2022). Reconceptualising regional inequalities: The application of a modified Philippe Martin model to the tourism sector. *Periodica Polytechnica Social and Management Sciences*. 31(1). DOI: https://doi.org/10.3311/PPso.17329

European Union (2024) NextGenerationEU, URL: https://next-generation-eu.europa.eu/recovery-and-resilience-facility_en

European Union (2025) Italy's recovery and resilience plan. URL: https://reforms-investments.ec.europa.eu/italys-recovery-and-resilience-plan_en#annual-events

Eurostat (2024a). Enterprises by NUTS 1 region and NACE Rev. 2 activity. Eurostat Database, European Commission. URL: https://ec.europa.eu/eurostat/databrowser/view/sbs_r_nuts2021/default/table?lang=EN

Eurostat (2024b). GERD (Gross domestic expenditure on R&D) by sector of performance and NUTS 1 region. Eurostat Database, European Commission. URL: https://ec.europa.eu/eurostat/databrowser/view/rd_e_gerdtot/default/table?lang=en

Eurostat (2024c). Gross domestic product (GDP) at current market prices by NUTS 1 region. Eurostat Database, European Commission. URL: https://ec.europa.eu/eurostat/databrowser/view/TEC00001/default/table

Fadeev, A., Alhusseini, S. (2023). Determination of urban public transport demand by processing electronic travel ticket data. *Periodica Polytechnica Transportation Engineering*, 51(4), 394–408. DOI: https://doi.org/10.3311/PPtr.21447

Fujita, M., Krugman, P., Venables, A. J. (1999). *The Spatial Economy: Cities, Regions and International Trade*. MIT Press. DOI: https://doi.org/10.7551/mitpress/6389.001.0001

Fujita, M., Thisse, J. F. (2002). *Economics of Agglomeration: Cities, Industrial Location and Regional Growth*. Cambridge University Press. 1–22. DOI: https://doi.org/10.1017/CBO9780511805660.001

ISTAT (2024). *Italy in Figures 2024.* Italian National Institute of Statistics, Rome. URL: https://esploradati.istat.it/databrowser/#/en/dw/categories

Jaber, A., Hamadneh, J. (2024). Demographic analysis of active transport mode users in urban context. *Periodica Polytechnica Transportation Engineering*, 52(4), 326–332. DOI: https://doi.org/10.3311/PPtr.23401

Jaffe, A. B. (1989). Real effects of academic research. *American Economic Review*. 79(5), 957–970. URL: https://www.jstor.org/stable/1831431

Juhász, M., Mátrai, T. (2024). Comprehensive sensitivity analysis of cost-benefit analysis variables for transport interventions. *Promet – Traffic and Transportation.* 36(4). DOI: https://doi.org/10.7307/ptt.v36i4.529

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review*. 70(5), 950–959. URL: https://www.jstor.org/stable/1805774

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*. 99(3), 483–499. DOI: https://doi.org/10.1086/261763

Krugman, P. (1993). On the relationship between trade theory and location theory. *Review of International Economics*. 1(2), 110–122. DOI: https://doi.org/10.1111/j.1467-9396.1993.tb00009.x

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1), 3–42. DOI: https://doi.org/10.1016/0304-3932(88)90168-7

Macsinka, K. (2025). Mobility impacts of urban development in the agglomeration. *Acta Technica Jaurinensis*. 18(4). DOI: https://doi.org/10.14513/actatechjaur.00782

Martin, P. (1998). Can regional policies affect growth and geography in Europe? *The World Economy*. 21(6), 757–774. DOI: https://doi.org/10.1111/1467-9701.00162

Martin, P. (1999a). Are European regional policies delivering? Sciences Po Economics Publications Working Paper. HAL Open Archive. URL: https://sciencespo.hal.science/hal-01011168

Martin, P. (1999b). Public policies, regional inequalities and growth. *Journal of Public Economics*. 73(1), 85–105. DOI: https://doi.org/10.1016/S0047-2727(98)00110-8

Martin, P., Mayer, T., Mayneris, F. (2011). Spatial concentration and plant-level productivity in France. *Journal of Urban Economics*. 69(2), 182–195. DOI: https://doi.org/10.1016/j.jue.2010.09.002

Martin, P., Ottaviano, G. I. P. (1999). Growing locations: Industry location in a model of endogenous growth. *European Economic Review*. 43(2), 281–302. DOI: https://doi.org/10.1016/S0014-2921(98)00031-2

Ötvös, V. (2025). The two sides of the coin of externality in road transport. *Acta Technica Jaurinensis*. 18(3). DOI: http://doi.org/10.14513/actatechjaur.00773

Pastuszka, S. (2017). Regional differentiation of the demographic potential in Italy and Poland. Comparative Economic Research. Central and Eastern Europe, 20(3), 137-159. DOI: https://doi.org/10.1515/cer-2017-0024

Petrović, N., Jovanović, V., Marković, S., Marinković, D., Nikolić, B. (2025). Multi-criteria decision-making approach for choosing e-bus for urban public transport in the city of Niš.*Acta Technica Jaurinensis*, 18(1). DOI: https://doi.org/10.14513/actatechjaur.00758

Prusov, D., Dubova, S. (2022). Modeling of passenger flows in agglomerations of large developing cities (on the example of Kyiv). *Periodica Polytechnica Transportation Engineering*. 5(4). DOI: https://doi.org/10.3311/PPtr.19785

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*. 94(5), 1002–1037. DOI: https://doi.org/10.1086/261420

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*. 98(5), S71–S102. DOI: https://doi.org/10.1086/261725

Sahin, S. N. A. (2023). A systematic overview of decision-making on public transport systems using bibliometric analysis on VOSviewer. *Transport and Telecommunication Journal*, 24(4). DOI: https://doi.org/10.2478/ttj-2023-0036

Škorupa, M., Bulková, Z., Gašparík, J., Kendra, M. (2024). Methodology for designing a system of public passenger transport in a functional region. *Periodica Polytechnica Transportation Engineering*. 52(1). DOI: https://doi.org/10.3311/PPtr.23159

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*. 70(1), 65–94. DOI: https://doi.org/10.2307/1884513

Szalmáné Csete, M., Baranyi, T. (2023). Can smart city planning enhance the sustainable transition of the EU capitals? A project and strategy-based smart, sustainable performance analysis in the programming period 2014–2020. *Production Engineering Archives*. 29(4). DOI: https://doi.org/10.30657/pea.2023.29.51

Szendi, D. (2024). Testing the validity of spatial beta convergence for the countries of Southern Europe. *Review of Business & Management – Theory, Management, Practice (TMP)*. 20(2), 55–67. DOI: https://doi.org/10.18096/TMP.2024.02.04

Van Leeuwen, E. (2015). Urban-rural synergies: An explorative study at the NUTS3 level. *Applied Spatial Analysis and Policy*. 8(3), 273–289. DOI: https://doi.org/10.1007/s12061-015-9167-x

Wang, Q., Yu, W., Wang, W., Hua, X. (2022). Towards intercity mobility system: Insights into the spatial interaction gravity model and determination approach. Promet-Traffic & Transportation. DOI: https://doi.org/10.7307/ptt.v36i2.414

World Bank (2024). World Development Indicators: Italy. World Bank Database. URL: https://data.worldbank.org/country/italy

Zheng, Z., Liu, Y., Zhou, Y., & Chen, T. (2026). Navigating from Spatial Heterogeneity to Synergistic Convergence: The Dynamics of Innovation Efficiency in China's Regional Innovation Ecosystems. International Review of Economics & Finance, 104933. DOI: https://doi.org/10.1016/j.iref.2026.104933

Zöldy, M. (2024). Changes at mobility space use in the cognitive mobility era. *Acta Technica Jaurinensis*, 17(4). DOI: https://doi.org/10.14513/actatechjaur.00750

https://doi.org/10.55343/CogSust.22150

Politics of knowledge: A questionable legal innovation

Stephan Kühnel
HTW Dresden
Dresden, Germany
Stephan.Kuehnel@htw-dresden.de

Abstract
German legal practitioners have tried to address complaints about professional exams. Two Supreme Court rulings suggest that a prediction of a candidate's future success in a professional examination cannot be verified or rebutted by a court or anyone else, but only by the individual examiner. In contrast to the apparent understanding of legal professionals, this type of knowledge is not an "unknown territory"; it is traditionally covered in research, e.g., in psychology or human resource management. It also undermines internationally recognised principles of sound assessment governance, such as validity, reliability, transparency, and fairness, as emphasised in UNESCO's global guidance on learning assessments and in OECD frameworks for evaluation and quality assurance in education systems. Situating these issues within a broader sustainability perspective, this paper argues that assessment regimes form part of society's knowledge infrastructure, and their integrity is essential for both cognitive sustainability – the long-term capacity of institutions to maintain transparent, evidence-based and error-correcting knowledge practices – and for the realisation of Sustainable Development Goals. In particular, SDG 4 (Quality Education) requires equitable, valid and reliable assessment mechanisms as a foundation of fair and merit-based professional pathways. At the same time, SDG 16 (Peace, Justice and Strong Institutions) calls for accountable, transparent and trustworthy institutional processes in all sectors, including regulated professions. Assessment systems that shield examiner judgment from external scrutiny impede these sustainability commitments by eroding public trust, weakening institutional resilience, and preventing the correction of systematic errors. This paper draws on a range of disciplines to highlight the decisions and errors made in practice. In addition, an International Education standard by the International Federation of Accountants (IFAC) is used to contrast findings. The paper thereby contributes to cognitive sustainability by examining how assessment systems can either sustain or undermine the long-term resilience and transparency of societal knowledge practices.

Keywords
Knowledge, Professional Exam, German Supreme Court Rulings

1. Introduction

As professional examinations are gatekeeping infrastructures that shape human capital formation and public trust, their design is entwined with sustainable development in two ways: first, as part of SDG 4 commitments to equitable, valid and reliable assessment; second, as elements of SDG 16 commitments to effective, accountable and transparent institutions. The present jurisprudence is compromised by institutionalising a non-verifiable form of judgment that blocks error-correction and undermines cognitive sustainability across the professions. This paper addresses two rulings of the German Supreme Court (BVerfG) on professional examinations. The decisions are based on one false assumption. The result is that errors will go uncorrected, and even blunt manipulation by examiners is possible; i.e., there is no quality control.

From a cognitive sustainability perspective, professional examination systems play a foundational role in maintaining the stability and adaptability of societal knowledge. When such systems prevent error-correction or external scrutiny, they compromise the long-term resilience of institutional learning processes and contribute to cognitive lock-in. Integrating insights from cognitive sustainability thus enables a deeper analysis of how legal assessment doctrines shape the durability, transparency and adaptability of professional knowledge infrastructures.

2. Rulings and their reflection in the legal literature

There are two rulings of the German Supreme Court dated 17 April 1991, which are considered the beginning of a "new era" (Zimmerling and Brehm, 2007: 282) in the subject matter. It should be noted that German law provides

https://doi.org/10.55343/CogSust.22150

little guidance on the content of decisions on whether a candidate passes a professional exam. Most legal principles have been developed through court decisions.

The first ruling of the Supreme Court (BVerfGE 84, 34) covered in this paper concerns admission to the legal profession (German: *zweites juristisches Staatsexamen*). The candidate argued that the examiners' evaluation was flawed. The court referred to the equity principle codified in the German Constitution (Art. 3 Sect. 1, GG), i.e., that a court's evaluation of the candidate would put the candidate at an advantage over other candidates who did not complain. There was no reference to the concept of error whatsoever. How did the judges come to their conclusion? They separated two things: (i) aspects that may be proven to be right or wrong using objective criteria, i.e. being based on established knowledge or practice, (ii) aspects that relate to the evaluation of the performance of a candidate (German: *prüfungsspezifische Wertung*). The latter aspect has been justified as follows: (a) personal experiences and expectations of the individual examiner form the basis of the decision on the performance of the candidate in a professional exam, (b) the decision is based on complex deliberations of the examiner that may not be captured in rules. It follows logically that the latter aspects cannot be proven right or wrong. The Supreme Court did not provide any evidence to support its view. The bottom line is that aspect (i) can be challenged by a court, while a court is forbidden to take action related to aspect (ii).

The second ruling of the Supreme Court (BVerfGE 84, 59) relates to a candidate for the medical profession. The reasoning was essentially the same as described above (BVerfGE 84, 59; Tz. 62 ff.); the context is the use of a multiple-choice format in the examination (BVerfGE 84, 59; Tz. 65).

When looking at more current legal literature (Fischer et al., 2026), not much has changed. The evaluation of a candidate's performance by the examiner shall be based on the examiner's personal assessments and experiences. It is asserted that a review of the performance of an examiner by a third party is neither possible nor meaningful (Fischer et al., 2026: 356, ref. 635). Again, no evidence to support this view is presented.

It should be emphasised that the concept of "error" has disappeared in this legal reasoning. The opinion of the individual examiner is the ultimate decision authority in a German professional exam that cannot be challenged. A court is explicitly prohibited from doing so (cf. reference to the German constitution in the Supreme Court's ruling mentioned above).

3. Critique from legal literature

The German Supreme Court has managed to divide the decision of passing a professional exam into two parts: One part (see aspect (i), mentioned above) may be challenged by a court, the other part (see aspect (ii), mentioned above) not. Consequently, an examiner has a strong incentive to avoid any statement where a court may rule against the examiner and focus on the other party instead (Zimmerling and Brehm, 2007: 581). Put differently, if an examiner says "this is wrong" when referring to a candidate's statement, a court can challenge his opinion. He may choose to say, "Given my experience, I am not convinced," or "This is not practically relevant" instead. This problem – let us call it plain and simple manipulation – is referred to as the issue of the smart examiner (German: *geschickter Prüfer*; Zimmerling and Brehm, 2007: 581).

4. Empirical evidence

The German Supreme Court's assumption that the examiner's evaluation cannot be wrong is easily refuted. Of course, there are errors in human judgment. They have been extensively researched in organisational science or psychology, to name just two domains. To illustrate, some effects shall be described as follows (selected from and based on Scholz et al, 2003):

https://doi.org/10.55343/CogSust.22150

Table 1: Selected empirical evidence on judgement errors, based on Scholz et al (2003)

Effect	**Description**	**Author**
Positive illusion	Overestimation of own skills, e.g., unrealistic self-perception, excessive optimism, illusion of control, etc. *Transferred to the context of this paper:* The examiner relies on his judgment even if he has no or limited experience in the specific area of expertise.	Taylor (1989)
Overconfidence	Trust in one's own estimates that is not warranted by facts *Transferred to the context of this paper:* The relevance of an error or debatable statement of the candidate may be exaggerated.	Kahneman and Tversky (1972)
Hindsight bias	An incorrect judgment is re-interpreted when new information appears. *Transferred to the context of this paper:* A failure of the examiner to learn from their own errors.	Pohl (1992)
Following the herd	If others agree, it must be right. *Transferred to the context of this paper:* When a member of a group with diverse areas of expertise witnesses another member's evaluation, that member tends to agree.	De Bondt and Thaler (1987)
Base rate fallacy	The probability of events is not reflected; typically, rare events are over-emphasised. *Transferred to the context of this paper:* The examiner is relying on odd or special cases rather than a realistic picture of the work in the profession.	Kahnemann and Tversky (1972)
Anchoring	The beginning of a deliberation determines the path or further steps. *Transferred to the context of this paper:* The examiner has a positive or negative impression of the performance of the candidate, which is not reviewed later in the process	Chapman/Johnson (2002)

This paper will not cover group decisions in detail. It should be noted that errors in group decision-making have been documented (e.g., Antoni, 2003). The legal literature covers group decisions that, from the perspective of a judge, are not fully comprehensible (cf. Fischer et al., 2026: 355, 567). As the decision on a candidate's performance is not the consensus of a group of examiners, but rather a judgment of truth, this topic shall not be covered in detail either.

A conceptual problem arises from the auditing literature. The public may have incorrect expectations regarding the performance of an audit (Expectation Gap, cf. Humphrey, 1997). Members of the group of examiners may be considered public, i.e., not members of the profession or not sufficiently knowledgeable in the domain in question. If there is scientific evidence that incorrect expectations are to be found, can a German Supreme Court assume that the expectations (or prejudices) of individual examiners are the ultimate authority for a decision? The answer is that the professional body administering the exam will have to operationalise the subject matter, expectations for a member of the profession.

https://doi.org/10.55343/CogSust.22150

5. Epistemology – or a critique of an opinion that there cannot be knowledge (ignorance)

The German Supreme Court has invented a type of knowledge that cannot be verified or refuted. It is assumed that this is a "space". Legal authors' reasoning (e.g. Fischer et al., 2026: 356) cannot be understood, so the impossibility of verifying the evaluation of an examiner is limited to the courts. Conversely, if legal scholars or practitioners were unable to do this while others can, a third-party expert on the subject matter would have to be included in the trial. I have not found any reference in the Supreme Court rulings or the legal literature mentioned above that implies that simply adding a third-party expert would solve the problem. Personal comment: It would be interesting to see a psychologist deciding who gets to practice law. I would assume this is not the result that the German Supreme Court has intended.

Epistemology is the theory of knowledge and forms part of philosophy. The traditional analysis of (factual) knowledge holds that three elements are required, resulting in a definition of knowledge as justified true belief:

a) justification, i.e. there are good reasons that support a belief (Bernecker and Dretske, 2000: 3);
b) truth, the assertion or statement is true (e.g. Horwich, 1998: 8 ff.); and
c) belief, i.e. it is held by someone (Moser et al., 1998: 28).

This knowledge concerns facts or *knowing that*. It should be emphasised that the requirement that knowledge must be true cannot be eliminated. Knowledge is conceptualized to be universal, so there shall not be an interpretation that "true" means "true for an individual" or "true for a group" (i.e. relativism, cf. Moser et al., 1998: 62). Considering that knowledge is about acquiring and teaching findings in sciences and humanities at universities worldwide, it is pretty obvious why such an understanding is insufficient. Others must be able to replicate results and confirm or disconfirm them. In short, a justified belief is not knowledge. It should be noted that this definition applies only to empirical knowledge. For example, mathematics would be considered non-empirical or a priori knowledge.

Furthermore, there is *know-how* (cf. Kühnel (2004). Ryle (1949) explained that "knowing how" (his phrase) is knowledge of how to perform activities well, i.e., correctly, efficiently, or successfully. There needs to be a standard for the performance of such an activity, which is used to guide an individual's action. Put differently, e.g., a clock also performs well if it displays the correct time; however, this is not considered knowledge of the clock. An individual must be able to identify and remedy errors, repeat an action, improve their performance, and profit from the example of others (Ryle, 1949: 28 f.). In order to do that, an individual must understand the activity, i.e. make assumptions about what to do (Ryle, 1949: 29). It must be noted that Ryle considers know-how as an integrated concept (one step), i.e. it cannot be replicated by knowing what to do and actually doing it (two steps).

Table 2: Simplified conceptualisation of the concept of knowledge in philosophy

Know that	**Know how**
Justified true belief	Perform an activity by applying standards of reference.

Making a large step toward the actual acquisition of knowledge, the concept of methodology, or the application of methods, shall be considered. In other words, holding a belief does not result from gut feeling. That a method must be applied to evaluate a candidate's performance escaped the German Supreme Court's attention. Moreover, of course, a method can be applied incorrectly. Borrowing a concept from a different but somewhat related domain – prediction of the future financial performance of a company (or individual) – there is the following distinction: (i) subjective evaluations represent the lowest level of assurance, (ii) quasi-objective (scoring) models allow for intersubjective validation and critique, while (iii) objective models are considered the most reliable type (cf. Baetge et al., 2004). Objective models are used, for example, to predict insolvency cases. Research is performed to identify factors that suggest a company's future breakdown.

It should be stressed that a judge's perspective may differ, as people may appear in court and make all kinds of assertions (whether true or not) that the judge must use as the basis for their ruling. From this perspective, the critical terms could be as follows:

https://doi.org/10.55343/CogSust.22150

Table 3: Simplified conceptualisation of knowledge in law

Know that	Know-how
Justified opinion	Convincing argumentation

* Note that "true" is not a required component of legal knowledge or a legal decision.

What the German Supreme Court effectively suggests is to add another type of knowledge to the list, i.e., a prediction of an individual's future performance in the profession (or, more generally, in their job). In other words, the above table would be modified as follows:

Table 4: Simplified conceptualisation of knowledge in law, extended with the two rulings of the German Supreme Court discussed in this paper

Know that	Know-How	Predict the future performance of an individual
Justified opinion	Convincing argumentation	• Not verifiable nor refutable, incl. no third-party expert opinion possible • Does not correspond to results from other disciplines

Although each discipline in the sciences or humanities may develop its own terms and concepts, the German Supreme Court has recognised a key point: "unable" or "cannot be proven right or wrong" are themselves testable propositions. Moreover, these propositions are highly problematic: (i) Knowledge that cannot be proven right or wrong does not fit the definition of knowledge in the philosophical traditional analysis. It may be a relativist concept of belief at best. (ii) That nobody – even a third-party expert – can verify the decision is not plausible. Professions are typically built around a common body of knowledge. (iii) There are academic disciplines that research the future performance of an individual (e.g. human resource management, psychology) or of an organisation (e.g. accountancy). In other words, this deliberation of the German Supreme Court does not withstand even a basic plausibility check.

6. Looking at the accountancy profession

The International Federation of Accountants (IFAC) has issued International Education Standards (IES) for its member organisations, i.e., professional bodies for accountancy in various countries (cf. IFAC, 2025: IES 6.2). In Germany, the respective professional organisation is the Wirtschaftsprüferkammer (WPK; see https://www.wpk.de/wpk/aufgaben/ as of 25. 3. 2025). IES define knowledge and skills (know-how) that a professional should possess. It is an authoritative source that should have been covered in legal interpretations (cf. Fischer et al., 2026).

Let us look at IES 6 "Initial Professional Development – Formal Assessment of Professional Competence" (Version as of March 2025). This standard formulates seven principles for a formal examination in the profession (IFAC, 2025: IES 6 BC.10). The new items on the list are *authenticity* and *integrity*.

IES 6.9: "IFAC member organisations shall be responsible for ensuring that the design, delivery, and oversight of assessment activities and processes to assess professional competence within professional accounting education programs formally have high levels of: Authenticity, Equity, Integrity, Reliability, Sufficiency, Transparency, Validity "

IES 6.A1 covers the concepts of knowledge: know that ("technical competence"), know how ("professional skills"), and adherence to professional principles ("professional values, ethics, and attitudes "). The latter emphasises only the two other types of knowledge, not a separate type from the perspective of philosophical traditional analysis. All of them can be tested – that is the idea of a professional exam.

This paper is not intended to analyse IES in detail, but certain aspects shall be highlighted. All principles in the list shall be used with equal prominence; i.e., a low level of performance (of the accountancy body) in one aspect may not be compensated for by a better result in another.

https://doi.org/10.55343/CogSust.22150

Table 5: Overview of principles for the formal assessment in the accountancy profession (IES 6)

Principle	Description	Reference
Authenticity	Assesses the intended learning outcomes in a way that reflects realistic situations that professional accountants may face	IES 6.A10-.A11
Equity	Fair and without bias, allowing all aspiring professional accountants an equal opportunity to complete the professional accounting education program	IES 6.A12-.A13
Integrity	Designed, delivered, and overseen to minimise the potential breaches of assessment security, improper administration and/or completion of an assessment	IES 6.A14-.A15
Reliability	Assessment activity consistently produces the same conclusion, given the same set of circumstances. According to IFAC, this is a principle based on public expectations, i.e. trust in the profession	IES 6.A16-.A17
Sufficiency	Evaluates the required professional competence with an appropriate balance of depth and breadth, knowledge, and application, and integration across a range of situations and contexts	IES 6.A18-.A19
Transparency	Details of an assessment activity, such as the competence areas and learning outcomes to be assessed, and the timing of the activity, are disclosed publicly.	IES 6.A20-.A21
Validity	Assesses the intended learning outcomes.	IES 6.A22-.A23

Let us compare the lines of argumentation of the German Supreme Court and IFAC on a selected set of aspects:

a) **Level of analysis**: The German Supreme Court does not distinguish between the examiner and the professional body, nor between the professional body's pronouncements. Errors, or even manipulation, by an individual examiner or a team of examiners will go unnoticed.

b) **Reliability and validity**: Not relevant, as the performance of the candidate and not the professional exam is in question. Methods applied are not covered, inadequate conceptualisations or errors in application will not be noticed.

c) Role of **judgement**: An aspect of the Reliability Principle is that a majority of assessors, acting independently, consistently come to the same judgment, given the same set of circumstances (IES 6.A16). Therefore, the professional body administering the examination will select appropriate assessors, provide them with an assessment rubric or marking guide, and provide training (IES 6.A17). The German Supreme Court does not even consider the possibility of an examiner's error. The individual examiner's personal experiences and expectations, and their complex deliberations that cannot be proven right or wrong, form the basis for a ruling on the subject matter.

7. Discussion

From a sustainability perspective, professional examinations are critical knowledge infrastructures that maintain the quality and social license of monopoly professions (law, medicine, accounting). If an assessment regime eschews validity, reliability, transparency, and equity as system properties – treating examiner judgement as non-refutable –, it erodes the long-term resilience of the professional ecosystem. This is misaligned with SDG 4 commitments (United Nations, 2015a) to quality and equity in education and assessment, and with SDG 16 (United Nations, 2015b) commitments to accountable, transparent institutions.

UNESCO's Education for Sustainable Development (ESD) agenda (UNESCO, 2021) explicitly links sustainable futures to how knowledge is assessed and governed, calling for assessment systems that are fit for purpose, technically rigorous, fair, and transparent. Treating assessment as a public good, with clear standards for validity, reliability, and moderation, is a prerequisite for sustainable knowledge transfer into the professions. OECD's reviews of evaluation and quality assurance likewise stress that credible systems must provide public accountability, align procedures and instruments with policy objectives, and support improvement through coherent assessment frameworks – all of which presuppose the contestability of judgement.

https://doi.org/10.55343/CogSust.22150

Framed as cognitive sustainability, these requirements can be stated as follows: institutions should conserve and renew their capacity for truth-tracking by maintaining open error-correction channels (external review, moderation, appeals), supporting epistemic diversity (drawing on psychology, HRM – Human Resource Management, educational measurement), and avoiding cognitive lock-in that freezes inferior practices. Recent work on cognitive sustainability and epistemic diversity in transitions provides conceptual scaffolding for such reforms. Finally, the UNESCO Futures of Education report and the OECD Learning Compass 2030 emphasise the societal need for adaptive, evidence-informed knowledge ecosystems – not merely curriculum goals, but governance practices that sustain learning, trust, and public value over time. Professional examinations that are auditable, transparent and method-guided are more consistent with these forward-looking frameworks than regimes that privatise judgment behind non-reviewable discretion.

Judges of the German Supreme Court have advanced an argument that must be considered a poor parody of basic scientific principles. Critical aspects are either defined away – e.g. the examiner cannot make errors or manipulate, the design of the assessment concerning the common body of knowledge of the profession is not considered at all – or presented in a light that almost certainly prevents the professional body from losing a court case – e.g. the reference to personal experiences and expectations of the individual examiner. The latter is presented without any reference to facts whatsoever.

In this paper, I have argued that the distinction between aspects that (i) may be right or wrong using objective criteria, i.e. being based on established knowledge and practice and (ii) an evaluation of the performance of a candidate based on personal experiences and expectations of the individual examiner and the complexity of a prediction of the future performance of a candidate (German: *Beurteilungsspielraum or prüfungsspezifische Wertung*) is flawed. The core of the German Supreme Court's ruling is that the latter cannot be verified by a court or by any third-party expert. When looking at philosophy or other disciplines, a type of empirical knowledge that nobody can verify or refute is not documented. On the contrary, researchers in other disciplines have developed methods to objectively evaluate an individual's performance (e.g., in human resource management or psychology). Even the simple fact that there may be judgment errors undermines the German Supreme Court's line of argument, as a court may no longer rely on the examiner's personal views as the ultimate decision criterion. It should be noted that the reliability of a professional examination that would be based on the views (and prejudices) of the individual examiner is effectively nil.

What to expect?

a. "Business as Usual"

The German Supreme Court has given professional organisations a toolkit that they or an individual examiner may use to manipulate the examination or hide errors. The success of this approach depends on whether candidates and the public accept the flawed line of reasoning described above.

b. German Professional Bodies

The problem described in this paper could become regarded as a worst-case scenario. The quality assurance system has broken down. The task of courts would have been to address errors in professional examinations, not to find an excuse for doing nothing. Candidates may reconsider their decision to join a profession. Customers may ask themselves why they should pay high fees for the professional service in question. The public may question its trust in the profession. Of course, there may be measures in a professional services company's quality management system that help mitigate some of the effects. However, this does not help the solo practitioner.

c. Administrative law

Cynically, the topic described in this paper could dramatically reduce the complexity of administrative law. Courts may look for comparable ways to redefine the problem so that public administration cannot make a mistake. This public administration would always be "right". Moreover, courts would have less work to do.

https://doi.org/10.55343/CogSust.22150

8. Passage of time and theoretical aspects towards a solution to the problem

This paper presents assertions made by the German Supreme Court in 1991. The question arises: how could these false assertions survive for so long?

Firstly, rulings of the German Supreme Court must be adhered to by other courts (§ 31 Sect. 1 BVerfGG). This is part of the law's systematic nature. A court ruling is presented as being related to discovering the truth. In the German Supreme Court, truth is referenced in the context of evaluating evidence. In German: *„Das Bundesverfassungsgericht erhebt den zur Erforschung der Wahrheit erforderlichen Beweis"*(§ 26 Sect. 1 BVerfGG). As discussed in this paper, the ruling itself is a justified belief and therefore not expected to be true. There are limits to challenging or changing a false ruling, although attempts are made to frame changes in jurisprudence systematically.

Secondly, the rulings discussed in this paper concern professions granted a monopoly. Without passing the professional exam, an individual may not practice in Germany. Some regulations are intended to protect the profession. For example, the concept of diligence relates to the professional conduct of auditors (§ 4 BS WP/vBP, in German: "Gewissenhaftigkeit "). § 4 Sec. 1 BS WP/vBP reads that an auditor is bound by law. In German: "*WP/vBP sind bei der Erfüllung ihrer Aufgaben an das Gesetz gebunden, haben sich über die für ihre Berufsausübung geltenden Bestimmungen zu unterrichten und diese und fachliche Regeln zu beachten*." WP is an auditor (German: Wirtschaftsprüfer), the typically used professional designation; vBP is a sworn auditor (German: vereidigter Buchprüfer) working only for small- and medium-sized clients. Access to the latter profession designation has been closed. Without explaining the meaning of this requirement in detail here, it should be obvious that any risk of non-compliance resides with the individual professional. It implies a risk of not being able to perform their job.

From an academic perspective, different approaches may offer concepts for addressing the issue highlighted in this paper. A preliminary selection is presented below.

IES 6 – covered in Section 6 of this paper – assumes that individual examiners may be erroneous. This raises the question of how the work of external experts – who did not participate in the decision on whether a candidate passed or failed the professional exam – would affect the exam's efficiency, transparency, and trustworthiness.[1]

At the organisational level, Cohen and Levinthal (1990) have proposed the concept of *absorptive capacity*. They have focused on a firm's innovative capabilities and defined the concept as the ability to recognise the value of new, external information, assimilate it, and apply it to commercial ends. Zahra and George (2002) have extended the concept by distinguishing between a firm's *potential and realised absorptive capacity.* They focus on the acquisition, assimilation, transformation, and exploitation of information. Their model puts activation triggers before the acquisition and assimilation elements. Mechanisms of social integration determine the transformation and exploitation elements (Zahra and George, 2002: 191 pp). Todorova and Durisin (2007) further elaborate on the concept by examining power relationships when exploiting information for competitive ends. This suggests future research could focus on activation triggers, social integration mechanisms, and power relationships in the German legal system. From an institutional perspective, the independence of individual judges is to be considered. In German: "Die Richter sind unabhängig und nur dem Gesetze unterworfen" (Art. 97 Abs. 1 GG).

Another theoretical approach concerns path dependence and lock-in. An organisation selecting a certain path enables future development but also inhibits alternative outcomes. This is because current knowledge is shaped by past experiences (Garud and Karnøe, 2001). Put differently, there is a link between mental models (ideas, beliefs) and institutions (Moreno-Casas, 2024). A comprehensive review by Goldstein et al. (2023) presents various cases of path-dependency and institutional lock-in and offers solutions. One important implication is to apply a forward-looking perspective, i.e. to focus on desired outcomes and then work backwards to break the lock-in. A useful synthesis is to construe the current doctrine as a form of cognitive path-dependence and institutional lock-in that suppresses absorptive capacity for external knowledge (psychometrics, educational measurement, organisational psychology). Cognitive sustainability – the durable stewardship of methods, standards and error-correction – requires activation triggers and integration mechanisms that keep institutions open to challenge and learning. The UNESCO and OECD (Taguma and Barrera, 2019). frameworks translate these into actionable design criteria (validity, reliability, moderation, transparency and accountability), while scholarship on epistemic diversity and justice cautions that excluding external expertise diminishes both legitimacy and problem-solving capacity.

[1] Thanks to one of the reviewers for suggesting to add corresponding thoughts.

https://doi.org/10.55343/CogSust.22150

9. Final comment

So where are we? The German Supreme Court has failed to distinguish between knowledge and opinion (see Table 4). Empirical knowledge is called empirical because it can be verified or refuted based on experience. German courts will have to find another excuse for non-action, or do their work.

Framing the findings through the lens of cognitive sustainability highlights that assessment systems are not merely administrative instruments but long-lived knowledge infrastructures whose design influences a society's epistemic resilience. Systems that suppress external review and deny the possibility of examiner error diminish institutional learning capacity and accelerate cognitive degradation. Sustainable, future-oriented governance therefore requires assessment regimes that maintain open pathways for challenge, validation, and continuous epistemic renewal.

From the academic perspective, selected theoretical concepts have been presented that may address the problem in the future. From a sustainability perspective, professional examinations are critical knowledge infrastructures that maintain the quality and social license of monopoly professions (law, medicine, accounting). If an assessment regime eschews validity, reliability, transparency, and equity as system properties – treating examiner judgement as non-refutable –, it erodes the long-term resilience of the professional ecosystem. This is misaligned with SDG 4 commitments to quality and equity in education and assessment, as well as with SDG 16 commitments to accountable, transparent institutions. UNESCO's Education for Sustainable Development (ESD) agenda explicitly links sustainable futures to how knowledge is assessed and governed, calling for assessment systems that are fit for purpose, technically rigorous, fair and transparent. Treating assessment as a public good, with clear standards for validity, reliability, and moderation, is a prerequisite for sustainable knowledge transfer into the professions. OECD's reviews of evaluation and quality assurance likewise stress that credible systems must provide public accountability, align procedures and instruments with policy objectives, and support improvement through coherent assessment frameworks – all of which presuppose the contestability of judgement.

Legal references

BS WP/vBP – Satzung der Wirtschaftsprüferkammer über die Rechte und Pflichten bei der Ausübung der Berufe des Wirtschaftsprüfers und des vereidigten Buchprüfers (Berufssatzung für Wirtschaftsprüfer/vereidigte Buchprüfer – BS WP/vBP). Link: https://www.wpk.de/fileadmin/documents/Wissen/Rechtsvorschriften/BS-WPvBP.pdf (as of 19. 12. 2024)

BVerfGE 84, 34. URL: https://www.servat.unibe.ch/dfr/bv084034.html

BVerfGE 84, 59. URL: https://www.servat.unibe.ch/dfr/bv084059.html

BVerfGG – Gesetz über das Bundesverfassungsgericht, as of 22. 12. 2025

GG – Grundgesetz für die Bundesrepublik Deutschland, as of 22. 3. 2025

References

Antoni, C. (2003). Teamarbeit. In: Auhagen, A. E., Bierhoff, W. (Hrsg.). *Angewandte Sozialpsychologie – Das Praxishandbuch*. Beltz, Weinheim. 351–365.

Baetge, J., Kirsch, H.-J., Thiele, S. (2004). *Bilanzanalyse*. 2. Auflage. Düsseldorf, IDW-Verlag.

Bernecker, S., Dretske, F. (eds). 2000. *Knowledge. Readings in Contemporary Epistemology.* Oxford University Press, Oxford. 13–15.

Chapman, G. B., Johnson, E. J. (2002). Incorporating the Irrelevant: Anchors in Judgment of Beliefs and Value. In: Gilovich, T., Griffin, D., Kahneman, D. (eds). *Heuristics and Biases: The Psychology of Intuitive Judgement.* Cambridge University Press, Cambridge, UK. 120–138. DOI: https://doi.org/10.1017/CBO9780511808098.008

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. Administrative science quarterly, 35(1), 128-152. URL: https://josephmahoney.web.illinois.edu/BA545_Fall%202022/Cohen%20and%20Levinthal%20(1990).pdf

De Bondt, W. F. M., Thaler, R. H. (1987). Further evidence on investor overreaction and stock market seasonality. *The Journal of Finance*. 42(3), 557–581. DOI: https://doi.org/10.1111/j.1540-6261.1987.tb04569.x

Fischer, E, Jeremias, C., Dieterich, P. (2026). *Prüfungsrecht*, 9th edition. C. H. Beck, München.

Garud, R., Karnøe, P. (eds) (2001). *Path dependence and creation*. Erlbaum, Mahwah, NJ.

Goldstein, J. E., Neimark, B., Garvey, B., Phelps, J. (2023). Unlocking "lock-in" and path dependency: A review across disciplines and socio-environmental contexts. *World Development*. 161, 106116. DOI: https://doi.org/10.1016/j.worlddev.2022.106116

Horwich, Paul. (1998). *Truth*. 2nd edition. Oxford University Press, Oxford.

Humphrey, C. (1997). Debating Audit Expectations. In: Sherer, M., Turley, S. (eds). *Current Issues in Auditing*. Paul Chapman, London.

IFAC (2025). *International Education Standard 6.* IFAC (International Federation of Accountants). URL: https://www.ifac.org/publications/international-education-standard-6-initial-professional-development-formal-assessment-professional or https://ifacweb.blob.core.windows.net/publicfiles/2025-03/IFAC-International-Education-Standard-6-Assessment.pdf (as of March 2025)

Kahneman, D., Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*. 3(3). 430–454. DOI: https://doi.org/10.1016/0010-0285(72)90016-3

Intuitive Judgement. Cambridge University Press, Cambridge, UK. 686–715. DOI: https://doi.org/10.1017/CBO9780511808098.041

Kühnel, S. (2004). *Wissensorganisation in Professional Service Firms: Perspektiven aus der Wirtschaftsprüfung.* Dissertation, Universität St. Gallen.

Moreno-Casas, V. (2024). A coevolutionary approach to institutional lock-in. *Review of Evolutionary Political Economy*. 5. 495–509. DOI: https://doi.org/10.1007/s43253-024-00132-2

Moser, P. K., Mulder, D. H., Trout, J. D. (eds) (1998). *The theory of knowledge – A thematic introduction*. Oxford University Press, Oxford.

Pohl, R. F. (1992). Der Rückschau-Fehler: Systematische Verfälschung der Erinnerung bei Experten und Novizen. *Kognitionswissenschaft*. 3. 38–44.

Ryle, G. (1949). *Concept of mind.* Hutchinson, London. Reprinted in 1958.

Scholz, R. W., Mieg, H. A., Weber, O. (2003). Entscheidung. In: Auhagen, A. E., Bierhoff, W. (eds). *Angewandte Sozialpsychologie – Das Praxishandbuch.* Beltz, Weinheim. 208–212.

Taguma, M., Barrera, M. (2019). OECD future of education and skills 2030: Curriculum analysis. URL: https://observatorioescolas2030.pt/wp-content/uploads/2023/11/OECD-FES_Learning-Compass-2030_Concept-notes.pdf

Taylor, S. E. (1989). *Positive illusions: Creative self-deception and the healthy mind.* Basic Books, New York, NY.

Todorova, G., Durisin B. (2007). Absorptive Capacity: Valuing a Reconceptualization. *Academy of Management Review*. 32(3). 774–786. URL: https://www.jstor.org/stable/20159334

UNESCO, P. (2021). *Reimagining our futures together: A new social contract for education*. Paris. URL: https://www.unesco.org/en/articles/reimagining-our-futures-together-new-social-contract-education

United Nations (2015a) SDG 4. Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all. *UN DESA SDGs*. URL: https://sdgs.un.org/goals/goal4

United Nations (2015b) SDG 16. Goal 16: Promote just, peaceful and inclusive societies. *UN DESA SDGs*. URL: https://sdgs.un.org/goals/goal16

Zahra, S. A., George, G. (2002). Absorptive Capacity: A Review, Conceptualization, and Extension. *Academy of Management Review*. 27(2). 185–203. DOI: https://doi.org/10.2307/4134351

Zimmerling, W., Brehm, R. G. (2007). *Prüfungsrecht: Verfahren, Vermeidbare Fehler, Rechtsschutz*. 3. Auflage. Carl Heymanns, Köln.

https://doi.org/10.55343/CogSust.22409

Artificial Intelligence and Sustainability: A Conceptual Framework for System-Level Impact Assessment

Adrienn Csernovszky
Budapest University of Technology and Economics, Department of Environmental Economics and Sustainability
Budapest, Hungary
adrienn.csernovszky@gmail.com

Maria Szalmane Csete
0000-0001-7170-9402
Budapest University of Technology and Economics, Department of Environmental Economics and Sustainability
Budapest, Hungary
csete.maria@gtk.bme.hu

Abstract

Artificial intelligence (AI) is rapidly emerging as a general-purpose technology with far-reaching implications for sustainable development. While AI applications are increasingly deployed across sectors such as healthcare, energy systems, urban management, and education, their overall sustainability impacts remain uncertain and often contradictory. Existing research typically examines isolated effects of AI within individual sustainability pillars, limiting understanding of systemic interactions, feedback loops, and long-term consequences. This study introduces a conceptual-analytical framework designed to assess the sustainability impacts of artificial intelligence across environmental, economic, and social dimensions, and extends it with an additional individual-level pillar. The framework defines a set of AI Impact Groups (AIG) that translate technological capabilities into system-level functions, including perception, learning, strategic foresight, coordination, and risk detection. In addition, the model introduces key input parameters – AI intensity, adoption level, autonomy, quality of use, and system quality – that influence how AI capabilities translate into sustainability outcomes. By linking AI capabilities, system-level functions, and sustainability pillars, the proposed framework enables a more integrated assessment of both opportunities and risks associated with AI deployment. The model highlights how AI impacts propagate across domains and may generate both short-term benefits and long-term systemic risks, such as rebound effects, technological dependence, or skill erosion. The framework provides a foundation for future scenario analysis, sector-specific impact assessment, and interdisciplinary collaboration to understand and govern AI-driven sustainability transitions.

Keywords
Artificial intelligence, Sustainability, SDGs, Systems approach, AI governance

1. Introduction

This study describes an ongoing research project conducted at the Department of Environmental Economics and Sustainability, Budapest University of Technology and Economics (BME). It aims to provide a general conceptual framework for assessing the impacts of artificial intelligence applications on sustainability, with a particular focus on risks and opportunities. Artificial intelligence (AI) has emerged as a key driver of technological and economic transformation and is increasingly adopted across a wide range of sectors and domains. However, the sustainability implications of AI remain uncertain, as its effects are often contradictory, making it difficult to determine whether AI ultimately supports or undermines long-term sustainability objectives (Katper at al., 2026). Existing research typically assesses AI impacts in a fragmented manner, focusing on one or two sustainability pillars or addressing isolated effects and outcomes (Antwi, Agyapong, Owusu, 2026). Therefore, we introduce the core definitions of artificial intelligence and sustainability, examine their interrelationships, and establish the need for a new modelling approach.

1.1 Basic concepts

Intelligence is defined as “the ability to learn and perform suitable techniques to solve problems and achieve goals, appropriate to the context in an uncertain, ever-varying world” (Manning, 2020: 1). Alan Turing asked the first foundational

https://doi.org/10.55343/CogSust.22409

question "Can machines think?" in 1935, and the Turing Test became a standard to measure machine intelligence (Doğan, Doğan, 2026). John McCarthy introduced the term Artificial Intelligence in 1955, who defined it as "the science and engineering of making machines that can perform tasks that would normally require human intelligence" (EIMT, 2026). Early artificial intelligence (AI) applications primarily focused on symbolic and logical problem solving. A notable example is the Logic Theorist, developed by Alan Newell and Herbert A. Simon, which demonstrated that machines are capable not only of performing calculations but also of solving problems through logical reasoning (Velupillai, Kao, 2014). Subsequent advances in machine learning and deep learning fundamentally transformed how machines learn from data. The work of Yann LeCun, Geoffrey Hinton, and Yoshua Bengio revolutionised deep neural network training, laying the foundation for a wide range of contemporary AI applications, including robotics, speech recognition, voice assistants, and language models. As a result, AI has evolved into a widely adopted technology across multiple sectors (Browning, LeCun, 2023; Matsuo et al., 2022).

AI types can be classified in multiple frameworks, and the terminology often overlaps or evolves as AI technologies advance rapidly. The term *AI types* generally refers to categories of AI systems, which are distinguished based on their technological approach, high-level functional role, or theoretical capability level (Table 1):

Table 1. Classification of AI systems by technological approach, function, and capability level

Classification basis	AI types
Technological approach	Symbolic AI, Machine Learning (including Deep Learning), Hybrid AI
High-level functional role	Perception AI, Prediction AI, Decision / Optimisation AI, Generative AI, Interaction AI, Autonomous AI
Theoretical capability level	Artificial Narrow AI (ANI), Artificial General Intelligence (AGI), Artificial Superintelligence (ASI)

Conversely, *AI functional categories* refer to groups of algorithms and mathematical approaches that define the main functions performed by AI systems. These functional categories enable specific *AI capabilities*, which represent the underlying abilities of AI systems. *AI applications* refer to the concrete, real-world uses of these functions in specific domains. Table 2 provides examples for all three levels: AI functional categories, AI capabilities (OECD, 2025) and AI applications.

Table 2. Examples of functional categories, AI capabilities (OECD, 2025), and applications

AI functional categories	AI capability	AI applications
Natural language processing (text understanding, text generation, speech processing)	Language	Chatbots, machine translation, text summarisation, virtual assistants
Affective computing and human–AI interaction (emotion recognition, conversational interaction)	Social interaction	Social robots, sentiment-aware customer service bots, and AI tutoring systems
Decision support, predictive analytics, optimisation	Problem solving	Medical diagnostic support systems, predictive maintenance, logistics optimisation
Generative AI and content generation	Creativity	AI image generators, text generation tools, music generation, and design assistants

https://doi.org/10.55343/CogSust.22409

Uncertainty estimation and AI-assisted decision making	Metacognition and critical thinking	Fraud detection systems flagging uncertain cases, risk assessment tools
Knowledge representation, information retrieval, and recommendation systems	Knowledge, learning and memory	Search engines, enterprise knowledge assistants, recommender systems
Computer vision (image recognition, object detection, visual analysis)	Vision	Medical image analysis, facial recognition, automated quality inspection
Robotic manipulation and physical automation	Manipulation	Industrial robotic arms, warehouse picking robots, and automated assembly systems
Autonomous systems and robotic control	Robotic intelligence	Autonomous warehouse robots, delivery robots, and agricultural robots

Global trends in AI development have been examined in the AI Index Report, which analyses data from the period 2022–2025 (SU HAI, 2025). The report characterises AI as one of the most transformative technologies of the 21st century. AI performance has surpassed human baseline levels in several domains (Figure 1), contributing to the rapid diffusion of AI systems and reshaping the boundaries between human and machine tasks.

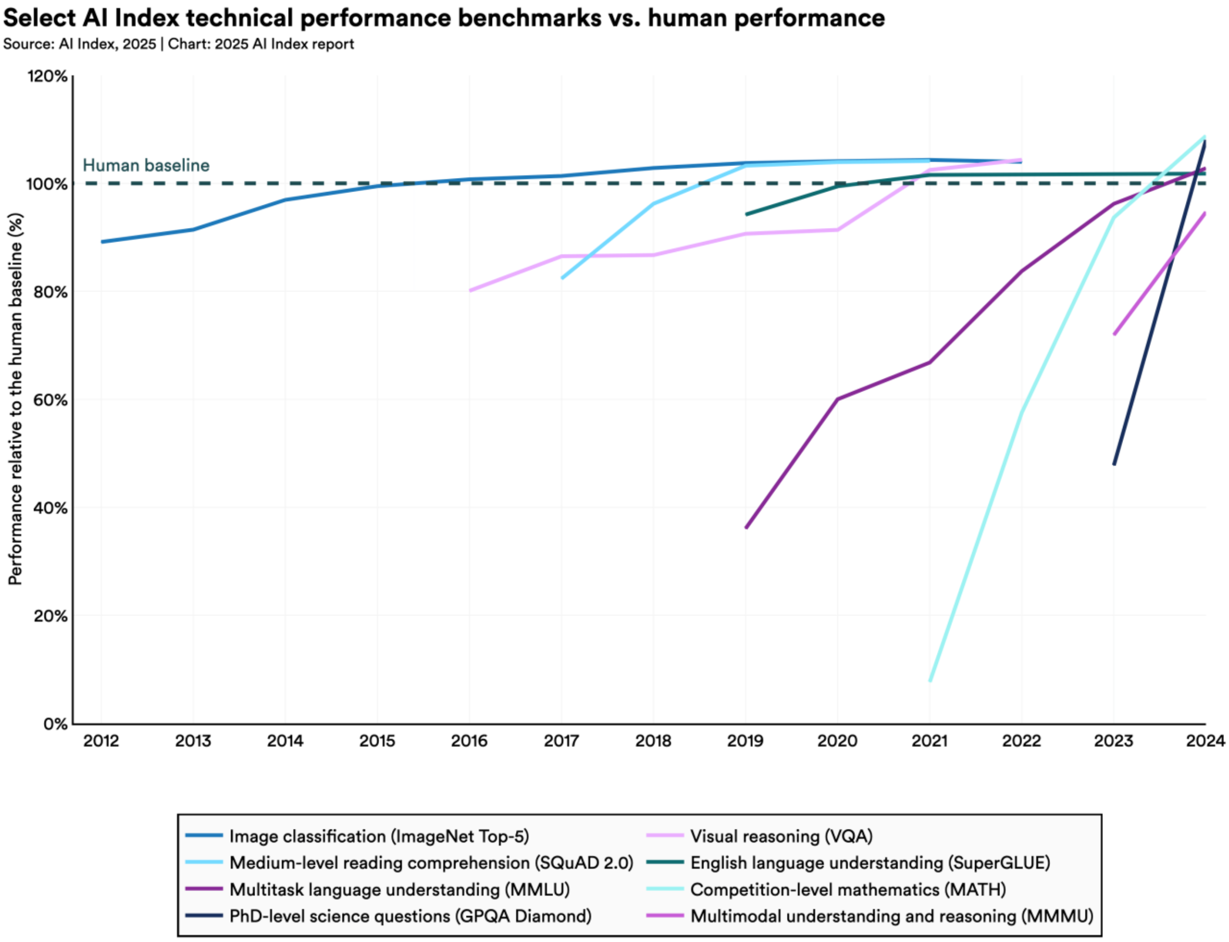

Figure 1. Performance relative to the human baseline (SU HAI, 2025)

https://doi.org/10.55343/CogSust.22409

In the past few years, AI has moved from the laboratory to daily life, with a strong presence in data-intensive sectors such as information and communication technologies, finance, healthcare, and manufacturing. The share of organisations applying AI in at least one function evolved from 20% in 2017 to 78% in 2024, based on 1363 participants representing a full range of regions (Singla et al., 2025).

AI optimism means that public perception of AI is positive and that it is seen as having more benefits than drawbacks. AI optimism shows a general upward trend across several countries, though in a few countries, stagnation or declining optimism can be observed. China has the highest level of AI optimism, with more than 80% of respondents agreeing, while the United States, the Netherlands, and Belgium report levels below 40%, based on data from 2022–2024. There are other perspectives to consider, as AI optimism is unevenly distributed and depends not only on technological performance but also on how visible and immediate the benefits of AI are, who experiences them directly, individual capabilities, and the surrounding AI regulations and policies. AI is mostly utilised in the economic and business domains, where productivity improvements enabled by AI are directly observable, which contributes to higher levels of AI optimism. In contrast, in sectors where tasks are easily automated, AI is associated with job insecurity, and among individuals with lower technical skills, it may be a source of perceived loss of control, leading to lower levels of AI optimism.

AI efficiency refers to how effectively AI applications convert consumed resources into achieved performance, and it is driven by fine-tuned small models and improved technologies and algorithms. Energy efficiency has improved by approximately 40%, while hardware costs have declined by around 30% annually. A rebound effect occurs, meaning that improvements in energy efficiency lower costs but also stimulate increased technology use. Consequently, AI efficiency gains do not necessarily lead to reductions in total resource or energy consumption. The rapid growth in data requirements necessitates more extensive data collection, storage, and processing, further intensifying the rebound effect. As AI has become a central concern due to its data-driven energy demand and societal impacts, governments have increased regulatory activity. In the United States, 59 AI-related regulations were issued in 2024 (SU HAI, 2025), and the European Union introduced the AI Act (Regulation (EU) 2024/1689).

The first definition of *sustainable development* was introduced in the Brundtland Report (Brundtland Commission, 1987) as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs." The United Nations' 2030 Agenda for Sustainable Development defines 17 Sustainable Development Goals (SDGs) addressing key global challenges such as climate change, equal opportunities, education, and healthcare (EurLex, n. d.).

1.2 Sustainability and AI

This section reviews how frequently AI tools are used in sustainability-related studies and identifies the domains in which they are most commonly applied. It also summarises prior research on AI's impacts on sustainability, including the main approaches and classification schemes. Finally, the key research gaps are highlighted, and it is described how the proposed methodology aims to complement and extend existing work by providing a different, more integrative perspective.

The number of AI applications in sustainability-related research has increased, as illustrated in Figure 2. AI is primarily used for forecasting, system optimisation, data mining, and remote sensing, as well as for accelerated experimentation and fast approximate simulation (Gohr et al., 2025).

https://doi.org/10.55343/CogSust.22409

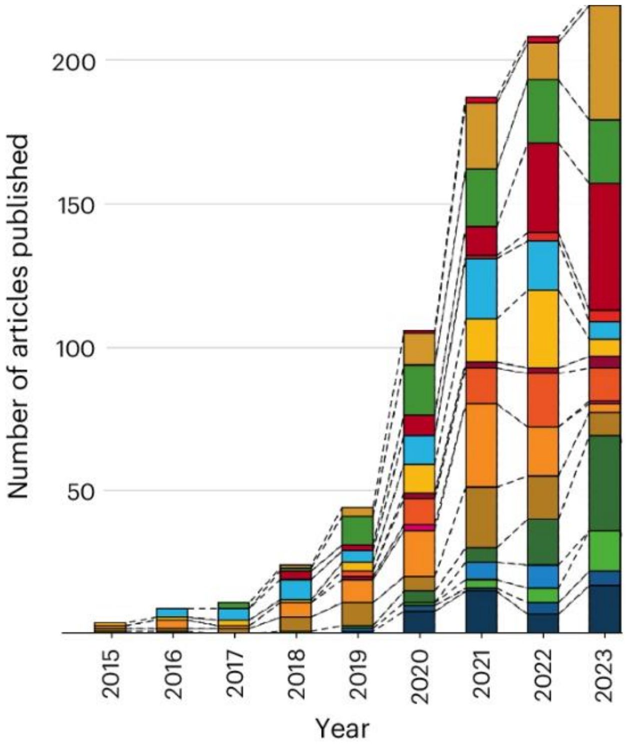

Figure 2. Number of AI applications in sustainability-related research (Gohr et al., 2025).

The same research classifies AI applications across sustainability topic groups and shows that the largest number of studies is in deep learning and supervised machine learning, particularly in applications involving system optimisation in water management and forecasting (Figure 3). The study reveals a methodological concentration, focusing on a limited number of sustainability domains or a narrow range of areas.

	Vegetation	Water	Forecasting	Remote sensing	Clean energy	Health care	Education	Industry	Total
Deep learning	22	44	32	26	22	8	9	2	165
Supervised machine learning	14	31	25	7	8	6	4	1	96
Evolutionary algorithms	5	11	10	5	13			2	46
Fuzzy logic	8	12	2		3				25
Intelligent decision support systems	5	11	2		3	2			23
Probabilistic reasoning	7	1	1	3	1	2			15
Computer vision	2			6	1	4			13
Symbolic AI	2	1	2	2	3	1			11
Optimization algorithms	1	5	1	2	1				10
Unsupervised machine learning	2		1	3		2		2	10
Multi-agent systems	2		1	4				1	8
Ensemble learning	2	1	3			1			7
Explainable AI			5		2				7
Generative AI				1		3	3		7
Natural language processing				1		4	2		7
Reinforcement learning			1	4				1	6
Large language models						2	3		5
Multi-criteria decision-making AI	1				2			1	4
Edge AI				3					3
Spatiotemporal AI			1		1				2
Total	73	117	87	67	60	35	21	10	470

Figure 3. Application frequency of the top 20 most frequent AI types across sustainability topic groups (Gohr et al., 2025)

https://doi.org/10.55343/CogSust.22409

Toderas (2025) analyses the complex role of AI related to sustainability. AI applications were highlighted for their positive impact, including optimising complex systems, monitoring across several fields, and smart city development, while challenges were primarily addressed through risk assessment and mitigation recommendations. The main risk categories were defined using a multi-pillar, functional methodology: ecological risks, socio-economic and ethical risks, and data and algorithm-related risks (Figure 4).

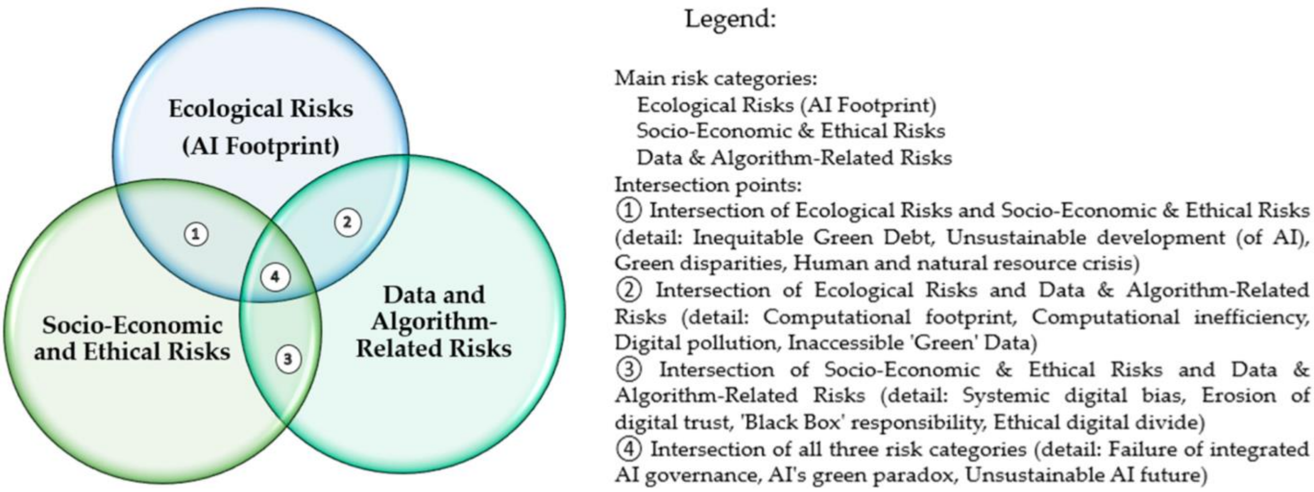

Figure 4. Main risk categories (Toderas, 2025)

The risk score was calculated by combining the probability of a risk's occurrence with its severity. Identified critical risks were (i) data bias, (ii) risk of amplifying inequalities and (iii) technological divide (AI divide) (Toderas, 2025).

The concept of the *AI green paradox* has emerged (Toderas, 2025). It describes a phenomenon in which AI applications intended to improve sustainability may increase overall energy use and emissions due to the hardware and software requirements of AI systems. This highlights the need to assess AI impacts across its entire lifecycle. A lack of transparency associated with so-called *black-box models*, as well as other socio-economic risks, has also been identified in relation to AI applications. *Green debt* refers to the risk that AI's short-term benefits are offset by its long-term impacts.

Mitigating these risks requires research in Green AI, the development of responsible governance frameworks, international collaboration, increased focus on education and training, and the development of "Green AI engineers" (Toderas, 2025). While this study adopts a multi-pillar perspective, it lacks an individual dimension of sustainability. Furthermore, there is a limited explanation of how impacts propagate across pillars, which input factors lead to specific outcomes, and how risks or opportunities may amplify over time.

1.2.1 AI impacts across sustainability pillars

Isolated effects, focusing on a single sustainability pillar or a single impact direction, despite the fact that AI and its sustainability consequences are complex and systemic. These impacts depend on sectors, regions, and time, and require analysing both benefits and risks, capturing interdependencies and cross-effects across sustainability pillars. Table 3 provides an overview of how AI-based applications generate both benefits and risks across environmental, economic, social, and individual pillars of sustainability, illustrating these dynamics with one example for each domain.

Table 3. Examples of AI applications and their associated opportunities and risks across sustainability pillars

Pillar	Domain	AI application examples	Potential sustainability benefits	Potential risks

https://doi.org/10.55343/CogSust.22409

Environmental	Resource and material use	Material and product design system with optimisation	Reduction of raw material use	More efficient and cheaper products may increase consumption (rebound effect)
Economic	Labour market and employment	Process automation system	Emergence of new sectors applying AI	Job displacement
Social	Equality	Digital educational assistant	Improved accessibility	Digital divide
Individual	Health and well-being	Mental health assistant	Health support	Excessive reliance on technology and risk of psychological dependency

AI impacts do not remain confined to a single pillar but often propagate across other pillars as well. For example, AI-optimised transportation can reduce emissions and improve air quality, leading to better public health outcomes and lower healthcare-related economic costs. This highlights the need to identify input factors that shape the direction of effects and to develop a common analytical framework capable of assessing all sustainability pillars.

1.2.2 AI applications across the UN Sustainable Development Goals

Artificial intelligence is increasingly recognised as a general-purpose technology with the potential to influence multiple dimensions of sustainable development. While the relationship between artificial intelligence and sustainability is often discussed in terms of environmental, economic, and social pillars, global policy frameworks typically structure sustainability challenges through the Sustainable Development Goals (SDGs) defined in the United Nations 2030 Agenda (UN, 2015).

Mapping AI applications across the SDGs provides a structured way to identify where artificial intelligence may help solve global challenges and where new risks may emerge. AI technologies are currently applied in areas such as healthcare diagnostics, energy system optimisation, urban management, climate modelling, and education. These applications demonstrate that AI can accelerate progress toward multiple SDGs simultaneously. However, the same technologies may also introduce systemic risks, including technological dependence, amplification of inequality, and increased energy demand associated with digital infrastructures.

Table 4 summarises selected examples of AI applications across key SDGs and highlights both potential opportunities and associated risks. The examples illustrate that AI-related sustainability impacts rarely occur in isolation and often involve complex interactions between environmental, economic, social, and individual dimensions. The examples presented in Table 4 demonstrate that artificial intelligence can support progress across multiple SDGs by improving efficiency, enabling better monitoring systems, and enhancing decision-making processes. In particular, AI has strong potential in areas characterised by complex systems and large datasets, such as healthcare diagnostics, energy systems, climate modelling, and urban management.

At the same time, AI-related sustainability impacts are inherently ambivalent. Technologies designed to improve efficiency may generate unintended systemic consequences. For example, improvements in energy efficiency enabled by AI may reduce operational costs, thereby stimulating greater use of digital technologies and contributing to rebound effects and higher overall energy consumption. Similarly, AI-driven automation may increase productivity and innovation while simultaneously contributing to labour market disruptions or new forms of technological dependency.

Another important observation is that AI impacts often extend beyond the domain in which the technology is initially deployed. For instance, AI-supported energy optimisation in urban infrastructure may reduce emissions (environmental pillar), improve economic efficiency (economic pillar), and enhance urban quality of life (social pillar). At the same time, increased reliance on AI systems may influence individual skills and competencies, particularly in decision-making, learning, and risk perception.

These examples highlight that the sustainability implications of artificial intelligence cannot be fully understood through isolated sectoral analyses. Instead, they require analytical approaches capable of capturing cross-domain interactions, feedback loops, and long-term systemic effects.

https://doi.org/10.55343/CogSust.22409

Table 4. AI applications across selected SDGs: potential sustainability benefits and risks

SDG	Domain	AI application examples	Potential sustainability benefits	Potential risks
SDG 3 – Good Health and Well-being	Healthcare systems	Medical image diagnostics, epidemic forecasting, and AI-based mental health assistants	Faster diagnosis, improved healthcare access, personalised treatment	Algorithmic bias, data privacy concerns, and overreliance on automated decision support
SDG 4 – Quality Education	Education systems	Adaptive learning platforms, AI tutoring systems, automated assessment	Personalised education, broader access to learning resources	Reduced critical thinking, digital divide
SDG 7 – Affordable and Clean Energy	Energy systems	Smart grid optimisation, renewable energy forecasting, and demand management systems	Increased energy efficiency, improved integration of renewable energy	Increased electricity demand of data centres, rebound effects
SDG 9 – Industry, Innovation and Infrastructure	Industrial production	Predictive maintenance, AI-driven product design, automated quality control	Higher productivity, resource efficiency, and accelerated innovation	Job displacement, technological dependency
SDG 11 – Sustainable Cities and Communities	Urban systems	Intelligent traffic management, smart waste management, urban climate modelling	Reduced congestion and emissions, improved urban services	Surveillance risks, data governance challenges
SDG 12 – Responsible Consumption and Production	Supply chains and the circular economy	AI-based supply chain monitoring, material optimisation, and waste sorting systems	Resource efficiency, waste reduction, improved transparency	Acceleration of consumption cycles, supply chain concentration
SDG 13 – Climate Action	Climate monitoring and modelling	Climate prediction models, carbon monitoring systems, and disaster forecasting	Improved climate risk management and mitigation planning	Carbon footprint of AI infrastructure
SDG 16 – Peace, Justice and Strong Institutions	Governance systems	AI-assisted public administration, corruption detection systems, policy simulation tools	Increased administrative efficiency and transparency	Algorithmic decision bias, reduced democratic oversight

The SDG-based overview shows that artificial intelligence applications simultaneously affect multiple sustainability domains and often create both opportunities and risks. However, these impacts within specific sectors or individual sustainability pillars limit the ability to analyse systemic interactions and cross-pillar effects.

https://doi.org/10.55343/CogSust.22409

To address this challenge, this study introduces a conceptual analytical framework designed to assess the sustainability impacts of artificial intelligence across interconnected dimensions. The proposed approach extends the traditional three-pillar sustainability model by incorporating an additional individual dimension. It introduces a set of AI impact groups and input parameters that enable a more systematic assessment of how AI capabilities translate into sustainability outcomes.

Figure 5 presents a general AI system evaluation framework covering the entire lifecycle. Evaluation is embedded across multiple phases, including design, data collection, model and system development, and deployment. Throughout these phases, various stakeholders conduct different forms of assessment – such as impact evaluation, data evaluation, model evaluation, model testing, benchmarking, and capability evaluation.

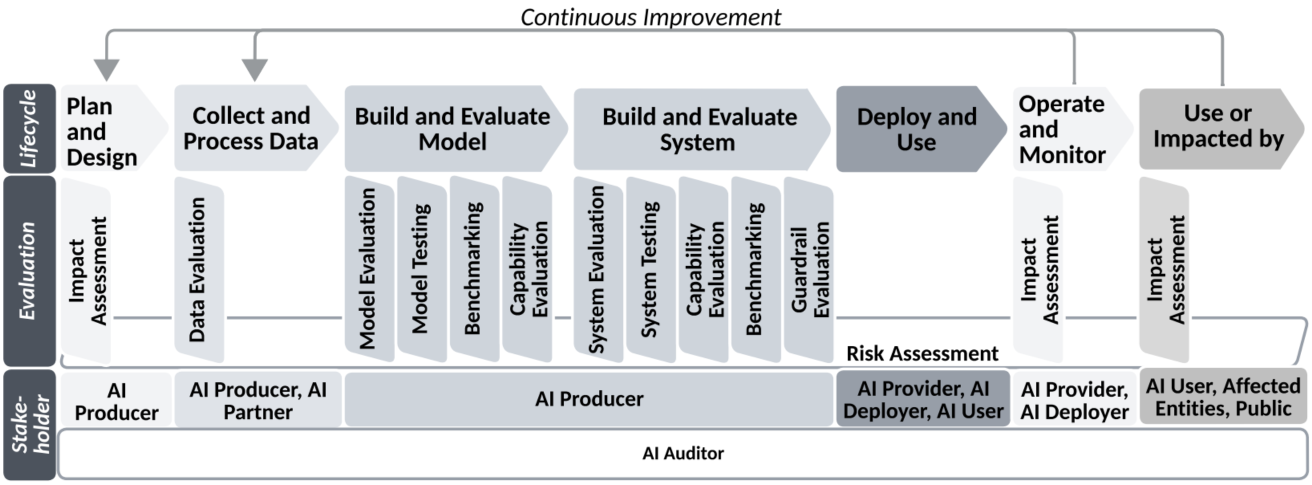

Figure 5. AI system evaluation from the point of view of supply chain (Xia et al., 2024: 78)

Humanity's last exam (HLE) (Phan et al., 2025a) provides an evaluation framework for large language models, consisting of 2500 challenging questions across a wide range of subject areas, contributed by nearly 1000 subject experts. Figure 6 compares the accuracy of several large language models across different benchmark types. While models achieve high scores on other benchmarks, HLE remains lower, indicating that it is more challenging and focuses on expert-level questions.

https://doi.org/10.55343/CogSust.22409

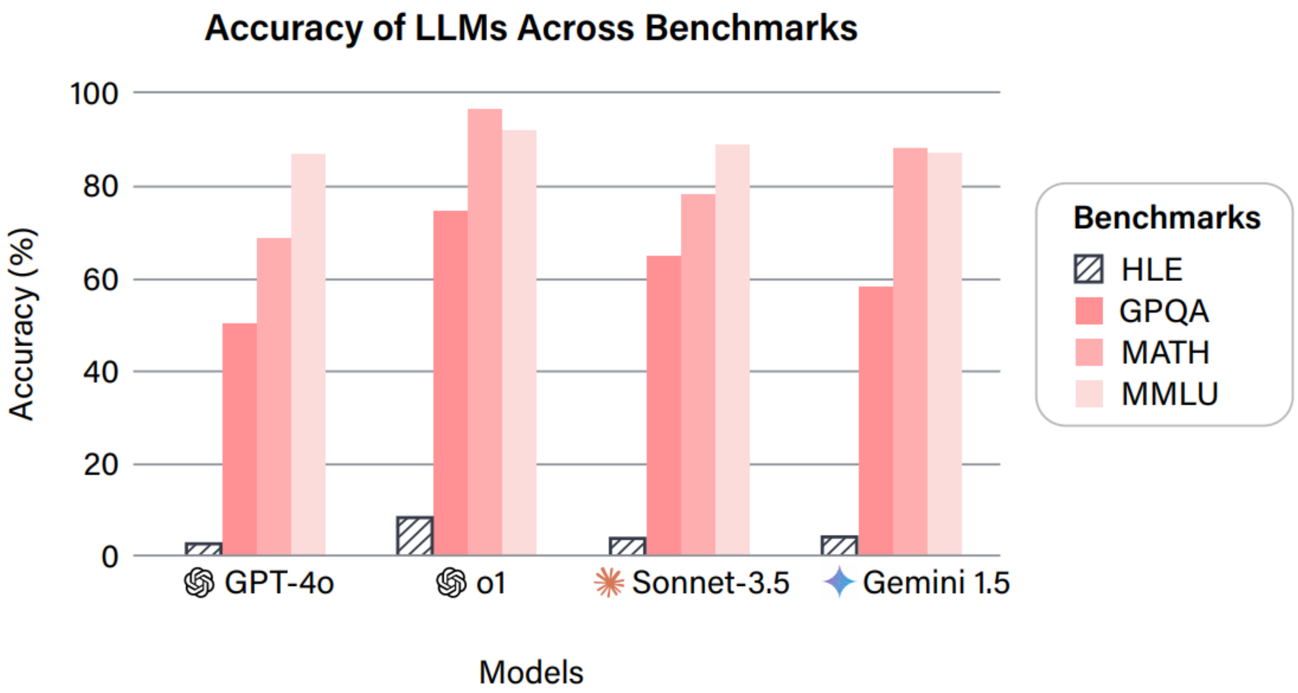

Figure 6. HLE and existing benchmark comparison in accuracy (Phan et al., 2025a)

Accuracy scores have increased over time: over 1.5–2 years, leading models improved from around 3–5% to 30–40% (see Figure 7). Competition among developers (OpenAI, Google, Anthropic, xAI) has been a key driver of this progress, contributing to higher accuracy levels. Newly introduced question sets and tasks are more informative for evaluation than repeated use of the same benchmarks, as models can be explicitly trained on previously released questions.

https://doi.org/10.55343/CogSust.22409

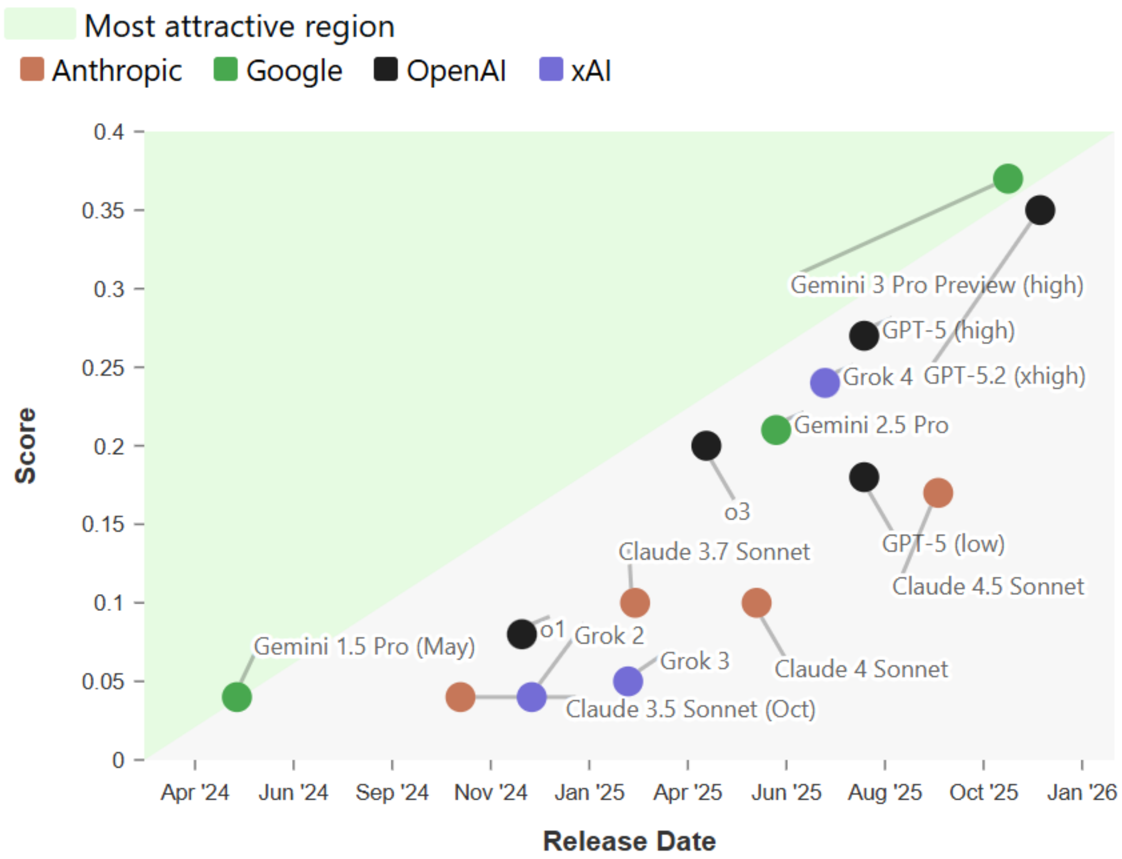

Figure 7. HLE benchmark leaderboard showing model accuracy scores over time. Results are based on evaluations reported by Artificial Analysis, an independent AI benchmarking and analysis company (Phan et al., 2025b)

Figure 8 shows the average number of completion tokens used across subject domains. Models tend to produce more tokens for physics, engineering and computer science tasks, whereas biology, medicine and humanities tasks require shorter responses. Differences in average token usage indicate that computational demands are not only model-specific but also domain-dependent. Longer reasoning chains involve more steps, increasing the likelihood of errors and making it more difficult to maintain coherence across extended responses. In such cases, the generation of intermediate reasoning steps combined with uneven or incomplete training data can further increase the risk of hallucinations and inaccuracies.

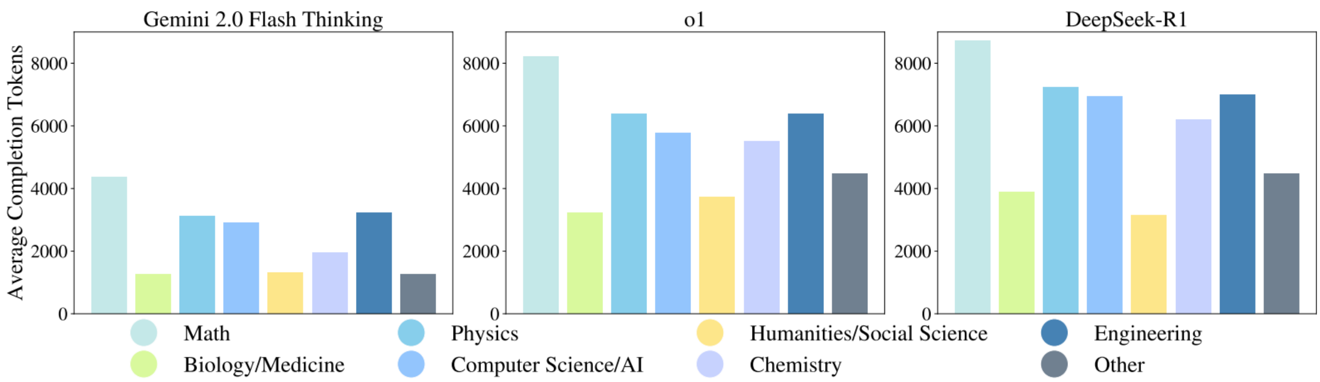

Figure 8. Average completion token counts across subject domains for selected AI models evaluated on the Humanity's Last Exam (HLE) (Phan et al., 2025a)

https://doi.org/10.55343/CogSust.22409

2. Framework description

This ongoing research is developing a new methodology for impact analysis. In this introductory study, the focus is on defining impact categories and input factors that will form the foundation of a model designed to assess impacts across sustainability pillars and the model's input factors.

2.1 Artificial Intelligence impact groups

AI impact groups are defined to enable the assessment of effects and the analysis of changes in skills, capabilities, and resilience. The proposed framework covers perception, interpretation, decision support, learning, risk and conflict detection, coordination, creative future-oriented thinking, motivation, and memory. Table 5 presents the AI impact groups. To emphasise the importance of each group, it also outlines the system-level consequences of their absence.

Table 5. AI impact groups and system-level consequences of their absence

No	AI Impact Groups - AIG	System-level consequences of absence
1	Strategic foresight and planning	The system is myopic, focusing only on current states and failing to account for delayed consequences.
2	Perception and pattern recognition	The system is blind, fails to detect warning signals and has delayed responses.
3	Knowledge integration and meaning formation	The system does not understand what it observes: the data are present, but lack coherence, and no interpretation is formed.
4	Learning and optimisation	The system does not optimise, repeat errors, and fails to adapt.
5	Risk and conflict detection	The system avoids difficult decisions.
6	Creative future construction and cyclical self-organisation	The system is unable to transform or develop; it optimises an outdated model and lacks alternative futures. Creative exhaustion.
7	Coordination and precision control	Intention does not translate into action; there is no execution.
8	Motivation, emotional and behavioural modelling	The system becomes exhausted, lacking persistence and commitment. Motivational collapse.
9	Memory and spatial-spatiotemporal modelling	The system does not store or properly utilise prior experience, resulting in short-term optimisation and decision-making.

Skill assessment example for pillars in case of 1. Strategic foresight and planning:

- Individual: system-level perspective, which helps to connect and integrate large-scale data patterns;
- Society: Societal system level perspective, which supports institutional transparency and strategic planning;
- Economy: Economic system level coherence, which covers a comprehensive assessment of economic performance and understanding of value change operations;
- Environment: Ecological system-level coherence, which helps in understanding the interdependencies of ecosystems and ecological networks.

2.2 Input parameters

2.2.1 Intensity

The depth of AI integration and the frequency of AI use are combined into a single input parameter, referred to as AI intensity. While depth of integration is primarily related to process, the frequency of use is more closely associated with individuals. Handling them separately would risk overrepresentation.

AI intensity is rated from 1 to 5:

1. Minimal integration and/or use (AI as experimentation):
   AI is rarely used in experimental settings or for isolated tasks. Without AI, there is no impact on system operation.
2. Low level of integration and/or use (AI as an optional tool):
   AI is used occasionally to support processes, but not the core of operations. Without AI, the system remains largely unchanged.

https://doi.org/10.55343/CogSust.22409

3. Moderate level of integration and/or use (AI as an integrated support layer):
   AI is widely used across many tasks, but it remains replaceable. Processes can function without AI, but less effectively. If AI fails, the system slows down or operates with reduced quality but continues to function.
4. High level of integration and/or use (AI has a mission-critical role):
   Processes are designed around AI, and it is embedded in core operations. If AI fails, the system is disrupted, and only partial operation is possible.
5. Full integration and/or high frequency of use (AI has a system-defining role):
   The system fully depends on AI; if AI fails, the system stops functioning.

Higher levels of AI intensity increase potential benefits while simultaneously amplifying systemic risks and dependencies. Meanwhile, lower levels of AI intensity are associated with limited impacts and competitive disadvantages due to missed efficiency gains.

2.2.2 AI adoption level

AI adoption level is defined as the extent to which AI applications are adopted within a sector.
AI adoption level is also rated from 1 to 5:

1. Very low adoption: AI is used in isolated experimental projects.
2. Low adoption: AI applications are used by a small number of actors within the sector.
3. Moderate adoption: AI applications are used by a significant share of actors within the sector.
4. High adoption: AI applications are used by the majority of actors in the sector.
5. Very high adoption: AI applications are used by nearly all actors within the sector.

AI adoption determines whether the impacts remain local or become systemic.

2.2.3 Autonomy

AI autonomy refers to the degree of autonomous operation of an AI application.
AI autonomy is rated from 1 to 5:

1. No autonomy (static AI): AI operates entirely under human control.
2. Low autonomy (assistant AI): AI supports tasks but executes decisions only with human approval.
3. Partial autonomy (supervised agent): AI operates under continuous human supervision.
4. High autonomy (autonomous operator): AI operates with human interventions only in exceptional cases.
5. Full autonomy (AI agent): AI operates without human intervention.

As AI autonomy increases, AI shifts from a supportive to a high-level decision-making tool. Reduced human oversight may amplify impacts and decrease the detectability of errors or unintended effects. Higher levels of autonomy may cover a broader range of tasks, potentially leading to skill erosion at the individual level, while also reducing human workload and improving operational efficiency.

https://doi.org/10.55343/CogSust.22409

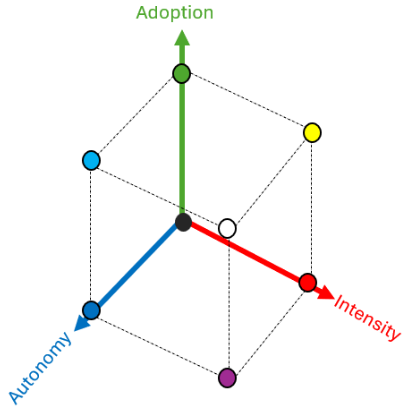

Figure 9. AI application space defined by adoption, autonomy and intensity

Figure 9 illustrates how the three input factors – AI adoption, autonomy and intensity – jointly define the space of existing AI applications, allowing any AI use case to be positioned within this framework. This input layer provides a functional classification, complemented by an additional layer focusing on quality, including quality of use and system quality.

2.2.4 Quality of use

Quality of use refers to the degree to which AI is applied consciously and critically, including how AI outputs are interpreted and evaluated. The following values can be assigned to this factor.

-1: Blind use: The output of an AI application is accepted without critical evaluation or validation.
0: Conscious use: The output of an AI application is accepted with some level of control, applying occasional validation.
1: Reflective use: The output of AI is interpreted and critically evaluated. AI supports learning, systemic understanding, and decision-making.

AI validation refers to the user confirming that the AI application is being used for its intended purpose. Quality of use is a key input factor as it shapes the direction of AI's influence on outcomes. Blind use may amplify the negative impacts of AI and contribute to unsustainable practices. In contrast, reflective use can lead to more informed, responsible and sustainable supportive decision making.

2.2.5 System quality

The last input factor relates to the evaluation of an AI application, including its robustness, performance, security and safety, reliability, and transparency. A wide range of AI application types exists, each with distinct evaluation processes. This section provides illustrative examples and proposes a unified rating approach.

In the present framework, the definition of the system quality input factor is kept general, as a wide variety of AI applications exist across sectors. Within a given domain, professionals with task-specific expertise are best positioned to evaluate the quality of an AI application on a scale of 1 to 5. To support this assessment, the following classification provides a structured reference framework:

1. Poor quality: Low data quality with frequent errors or hallucinations without any validation and risk management, and exhibiting a poor level of security.
2. Low quality: Occasional error handling combined with poor data quality.
3. Moderate quality: Acceptable performance with partial validation, basic monitoring, and risk management functions applied in several cases.
4. High quality: Strong data and model quality, regular validation, continuous monitoring and risk management application.
5. Very high quality: Verified and stable performance supported by high quality of data, robust and secure operation, documented processes and measurable compliance.

https://doi.org/10.55343/CogSust.22409

An evaluation of system quality highlights that the impacts of AI on sustainability also depend on model performance, deployment context, governance mechanisms, and continuous monitoring.

3 Discussion

The primary goal of this ongoing research is to develop an analytical framework for assessing the impacts of artificial intelligence across the sustainability pillars. This research investigates how the effects of AI applications propagate across the different pillars of sustainability, what determines whether these impacts are positive or negative, and which input factors shape the sustainability impacts of AI.

The analysis first reviews key concepts related to artificial intelligence, including AI types, functional categories, capabilities, and application examples, and discusses related phenomena such as efficiency gains, rebound effects, and the AI green paradox. Global AI trends show that artificial intelligence is becoming one of the most transformative technologies and is increasingly used in sustainability-related research.

Existing approaches focus on isolated effects, single domains, or specific AI applications; the proposed methodology captures cross-pillar interactions, feedback mechanisms, and the role of input factors. By extending the traditional three-pillar sustainability framework with an individual dimension, the model enables an assessment of how AI influences skills, capabilities, resilience, and long-term system dynamics. The framework provides a policy-relevant analytical approach by linking AI-driven changes to sustainability challenges addressed in the Sustainable Development Goals (SDGs), while allowing examination of how these changes may simultaneously influence progress toward multiple SDGs and potentially generate cross-domain trade-offs.

AI impact groups are defined in this study to translate AI capabilities into system-level functions. Examining system-level consequences in the absence of these impact groups indicates that performance metrics alone are insufficient and that AI effects should be interpreted as system-level phenomena. The proposed approach extends existing work by linking AI capabilities to sustainability-relevant system behaviour rather than isolated outcomes.

The following input parameters are proposed: AI intensity, adoption level, autonomy, quality of use, and system quality. Quality of use and quality of system emerge as critical factors, as the blind application of a low-quality system may amplify negative effects. These parameters are also relevant for governance and policy design, as they help identify conditions under which AI applications support sustainable development goals or generate unintended systemic risks. Input factors collectively shape AI's impacts and should therefore be assessed together.

Sustainability is commonly structured around three core pillars: environmental, economic, and social dimensions. In addition to these pillars, the individual can be conceptualised as a fourth, cross-cutting pillar embedded within and interacting with all three dimensions. The proposed model extends the traditional three-pillar sustainability framework by incorporating the individual as a fourth pillar, resulting in a 3+1 framework. The model links AI capabilities to changes at the skill and parameter levels, enabling a systematic analysis of cross-pillar interactions across time and regions. It can provide insights into how AI may simultaneously generate short-term benefits and long-term risks. AI may accelerate the development of fusion power plants and quantum computers, leading to a new technological paradigm shift that feeds back into AI and reshapes its broader systemic impacts. These complex feedback loops and inherent uncertainties underscore the need for a structured, shared analytical framework. The goal is to develop a methodology and an initial model that supports collaboration among researchers and experts within specific sectors or sustainability pillars.

4 Future research directions

Future research will focus on further developing the proposed model by refining input factors and systematically analysing their interactions and cross-effects. Beyond structural refinement, particular emphasis will be placed on integrating a cognitive sustainability perspective, examining how knowledge generation, perception, and decision-making processes shape and are shaped by the model's inputs and outputs. This includes exploring what external and internal conditions influence input parameters, identifying additional latent factors, and integrating regional and temporal dimensions to capture spatial variation and long-term dynamics. The model is also intended to support forward-looking analysis and scenario-based assessment, with a specific focus on enhancing cognitive accessibility and usability for diverse stakeholders. Further work will also address the selection and development of appropriate computational tools that allow domain experts to interact with the model, contribute assessments and collaboratively add use cases. Particular attention will be given to incorporating feedback loops and cross-pillar interactions. This may enable the model to represent how impacts propagate between sustainability pillars over time, while also reflecting how these dynamics are perceived, interpreted, and acted upon by decision-makers. In this context, the model aims to bridge analytical rigour with cognitive realism. Finally, future efforts will focus on developing effective visualisation approaches to support transparency, interpretability, and practical use of the model across research and policy contexts.

References

Antwi, B. O., Agyapong, D., & Owusu, D. (2026). Beyond adoption: Analysing the sufficient and necessary factors for AI-driven sustainable performance in an emerging economy. Sustainable Futures, 11, 101750. DOI: https://doi.org/10.1016/j.sftr.2026.101750

Browning, J., & LeCun, Y. (2023). Language, common sense, and the Winograd schema challenge. Artificial Intelligence, 325, 104031. DOI: https://doi.org/10.1016/j.artint.2023.104031

Brundtland Commission (1987). *Report of the World Commission on Environment and Development: Our Common Future*. United Nations. URL: http://www.un-documents.net/our-common-future.pdf

Doğan, A. R., & Doğan, A. İ. (2026). From the Turing Test to AI detectors: an epistemological mismatch in scholarly publishing. Brazilian Journal of Anesthesiology, 76(2), 844740. DOI: https://doi.org/10.1016/j.bjane.2026.844740

EIMT (2026). Who is the Father of AI and Machine Learning? *European Institute of Management & Technology*. URL: https://www.eimt.edu.eu/who-is-the-father-of-ai-and-machine-learning

EurLex (n. d.). Glossary of Summaries. Sustainable development. *Official website of the European Union*. URL: https://eur-lex.europa.eu/EN/legal-content/glossary/sustainable-development.html

Gohr, C., Rodríguez, G., Belomestnykh, S. et al. (2025). Artificial intelligence in sustainable development research. *Nature Sustainability*. 8, 970–978. DOI: https://doi.org/10.1038/s41893-025-01598-6

Katper, N., Hamid, A. B. B. A., Kaur, K., & Mohy-ud-Din, K. (2026). Dual pathways toward sustainability performance through AI technologies: Moderated mediation role of green HRM by fintech. Journal of Cleaner Production, 554, 148038. DOI: https://doi.org/10.1016/j.jclepro.2026.148038

Manning, P. C. (2020). Artificial Intelligence Definitions. Stanford University Human-Centered Artificial Intelligence. URL: https://hai-production.s3.amazonaws.com/files/2020-09/AI-Definitions-HAI.pdf

Matsuo, Y., LeCun, Y., Sahani, M., Precup, D., Silver, D., Sugiyama, M., ... & Morimoto, J. (2022). Deep learning, reinforcement learning, and world models. Neural Networks, 152, 267-275. DOI: https://doi.org/10.1016/j.neunet.2022.03.037

OECD (2025). Introducing the OECD AI Capability Indicators. OECD Publishing. DOI: https://doi.org/10.1787/be745f04-en

Phan, L., Gatti, A., Han, Z. et al. (2025a). Humanity's Last Exam. *arXiv*. DOI: https://doi.org/10.48550/arXiv.2501.14249

Phan, L., Gatti, A., Han, Z. et al. (2025b). Humanity's Last Exam Benchmark Leaderboard: Score vs. Release Date. *Center for AI Safety*. URL: https://artificialanalysis.ai/evaluations/humanitys-last-exam

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689

Singla, A., Sukharevsky, A., Yee, L., Chui, M., Hall, B. (2025). *The state of AI in early 2024*. Quantum Black, McKinsey. URL: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-2024

SU HAI (2025). The 2025 AI Index Report. *Stanford University, Human-Centered Artificial Intelligences (SU HAI).* URL: https://hai.stanford.edu/ai-index/2025-ai-index-report

Toderas, M. (2025). Artificial Intelligence for Sustainability: A Systematic Review and Critical Analysis of AI Applications, Challenges, and Future Directions. *Sustainability*. 17(17), 8049. DOI: https://doi.org/10.3390/su17178049

Velupillai, K. V., & Kao, Y. F. (2014). Computable and computational complexity theoretic bases for Herbert Simon's cognitive behavioral economics. Cognitive Systems Research, 29, 40-52. DOI: https://doi.org/10.1016/j.cogsys.2013.07.005

UN – United Nations (2015). Transforming our world: the 2030 Agenda for Sustainable Development. *UN Department of Economic and Social Affairs*. URL: https://sdgs.un.org/2030agenda

Xia, B. Lu, Q., Zhu, L., Xing, Z. (2024). An AI System Evaluation Framework for Advancing AI Safety: Terminology, Taxonomy, Lifecycle Mapping. *AIware 2024: Proceedings of the 1st ACM International Conference on AI-Powered Software*. 74–78. DOI: https://doi.org/10.1145/3664646.3664766