ercim-news.ercim.eu

Number 143 January 2026

ERCIM NEWS

Special theme:

AI for Science

Also in this issue's Research and Innovation section:

SmartCityHub, an Enabler for Trusted AI Solutions

Editorial Information

ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 2,000 printed copies and is also available online ,at https://ercim-news@ercim.eu.

ERCIM News is published by ERCIM EEIG
BP 93, F-06902 Sophia Antipolis Cedex, France
+33 4 9238 5010, contact@ercim.eu
Director: Dominique Hazaël-Massieux,
ISSN 0926-4981

Contributions
Contributions should be submitted to the local editor of your country

ERCIM News online edition
https://ercim-news.ercim.eu

Subscription
Subscribe to ERCIM News by sending an email to en-subscriptions@ercim.eu

Cover photo by Google DeepMind on Unsplash

JOINT ERCIM ACTIONS

SPECIAL THEME

NEXT ISSUE

ERCIM News 144, April 2026
Special theme: Open Science Experiences and Prospects

Inclusion is not Optional: Building Diverse and Sustainable Futures in Science

In this interview, Miriam Santos from the University of Porto and winner of the 2025 ERCIM Cor Baayen Award shares her perspective on why inclusion and diversity are essential for responsible science and technology. Drawing on her academic journey and her experience as founder of As Raparigas do Código, she reflects on the challenges of sustaining volunteer-driven initiatives, the lessons learned along the way, and the concrete actions research institutes and universities can take to foster more inclusive academic environments.

Figure 1: Raparigas do Código iprovide free education in technological fields to young girls.

Why do you think it is important to promote inclusion and diversity in research institutes, universities, and society at large?

The science and technology we're developing needs to be inclusive and diverse; otherwise it risks being biased and negatively affecting a significant amount of our society (women, ethnic minorities, people with disabilities, neurodivergents). Focusing on inclusion and diversity also allows us to counterbalance each other's blind spots when we are researching or building a new product, application, or system. I believe that this is how we will collectively "survive" the changes that are imposed on us daily, with the current advances of science.

You are the founder of the non-profit organisation As Raparigas do Código. Can you briefly explain its aim and how it operates?

As Raparigas do Código is a non-profit organisation that aims to provide free education in technological fields to young girls and women of all ages and backgrounds. Nowadays, technology is the main driver of financial stability, personal empowerment, opportunity, innovation, and social good, and we believe that education is the single most fundamental piece that we can provide, free of cost, to anyone. We started by organizing several short courses for women during the pandemic (web development, introduction to programming, data science, personal branding) and then we moved on to planning and structuring activities for kids (in-person events) focusing on computer science, cryptography, artificial intelligence, and many others.

What motivated you to start this organisation? And where did you start from?

As I started to "climb the academic ladder", I realised there were less and less women in my circle. I was one of five women in my Master's specialisation (Informatics), and the only one moving to a PhD in Intelligent Systems. There were very few female professors in Informatics Engineering (less than Physics and Mathematics), and that bothered me because I wanted to become one. The "invisibility" of it all was something that I felt I needed to change, and that perhaps I could help to change. So by 2020, I pitched the idea to my two (male and only) PhD colleagues at the time and we started a website [L1] offering free, online-courses and mentorship to women. In a short amount of time we grew a large community of women interested in learning more about technology and tech professionals (both men and women) wanting to join the team and become mentors.

How do you keep it alive? How did you managed to make it so successful?

It's hard. I feel we created the project at a golden time, because the pandemic pushed people to remain connected

Figure 2: Team members of Raparigas do Código, a non-profit organisation offering free technology education for girls and women.

somehow (it was "easier" to stay tuned to online lessons on saturday morning). At the time, the tech market also grew exponentially (there was a huge need for tech professionals, people could work from home), and companies were eager to invest to access diverse talent pools. It was also the beginning of a "general awareness wave" around the topics of gender equality, diversity, and inclusion, and we were "in the right place at the right time". Since then, five years have passed and it is hard to keep mentors motivated for so long. Some join and some leave, with all the dynamics this brings to the group (we are all volunteers, so we work outside working hours, in our free time). We keep it alive by making it innovative, fun, and people-centred. We eventually realised that motivation arises when we are also learning and doing something challenging, when we get to code fun projects and games [L2], and when we get to witness the impact we might have in others. So we associate often with other organisations to help them build solutions, organise events, and foster education, independently of "who gets credit" for it.

What are the main challenges that you have faced? Is there any "mistake" that is important to avoid?

Motivation for once (both mine and the team's) when committing to a project for so long, lack of experience running a team and an organisation in the beginning (with all the bureaucracy associated), lack of investment and support (funding to operationalise activities and events), burnout. One mistake I wish I had avoided was going "fast and furious" when starting the project. We grew quite fast, and it consumed some of us quickly. I believe it is similar to growing a startup sometimes: we agree to projects and ideas without the full capacity to deliver them, we want to say "yes" to everyone and everything, and we could not find the time to build a solid foundation to support that growth. Slow and steady wins the race, I guess.

How could research institutes and universities promote and support these initiatives?

Research institutes and universities are really in a position to drive change from within. Support women in their careers: promote growth opportunities for female academics, invest in women-led research projects and initiatives, create visibility for female academics within your departments, be mindful of gender representation when assembling juries, keynote speakers and chairs in conferences, and especially when distributing bureaucratic non-academic work which falls disproportionately onto women. It is also important to drive change early on: collaborate with schools in your district, foster open-days or summer schools led by some of your students, foster the creation of junior research labs for women (e.g., coding clubs initiatives), and maybe consider these types of initiatives as part of academic work (so that it can be included in the professional merit of professors or considered to reduce their schedule). Open up scholarships for promising female talent, award exceptional talent, adjust teaching and evaluation methods considering the historical and cultural biases and glass ceiling affecting women for decades. Promote a healthy, safe, and inclusive environment in your classes, especially if girls compose a minority of your class. There are so many ways in which to drive change.

Do you have any lesson learned or best practices that you would like to share with other researchers who are considering following a similar path?

Surround yourself by people who believe in you, who challenge you to be more, to be better, to aim higher. And by all means, do not let fear of failure or toxic voices stand in the way. Be unapologetically yourself, especially in academic environments that still reward conformity over impact: eventually your community will follow suit.

The interview was conducted by Monica Divitini, NTNU, chair of the ERCIM Human Capital Task Group.

Links:

[L1] https://raparigasdocodigo.pt

[L2] https://jogodasprofissoes.pt

Please contact:

Miriam Santos

University of Porto, Portugal

miriam.santos@fc.up.pt

Inclusive Digital Futures: Developing Technology and Culture - A new ERCIM Working Group

by Magdalini Chatzaki (FORTH-ICS)

ERCIM has launched a new Working Group on Inclusive Digital Futures: Developing Technology and Culture to embed inclusion at the core of digital research, education and innovation. It brings together researchers, administrators and leaders to drive systemic and lasting change across the ICT ecosystem.

As computer science shapes almost every dimension of contemporary life, questions of social responsibility have become central to technological progress. Digital systems are no longer neutral tools. They influence how societies function, how economies evolve and how cultures are expressed and transformed. This raises fundamental questions about who designs these systems, whom they serve and which values they encode.

The Inclusive Digital Futures Working Group responds by placing these questions at the heart of computer science, informatics, mathematics and ICT research and innovation. It starts from the premise that technological development and social progress are inseparable. If digital systems are to meet the needs of diverse societies, the ways in which they are conceived, built and governed must reflect that diversity.

Many ERCIM member institutions already host valuable initiatives in this area. Yet such efforts often rely on individual commitment and remain dispersed because of limited coordination and shared resources. Their long-term impact is therefore constrained. The Working Group seeks to overcome this fragmentation by promoting a more coherent and sustainable approach that integrates inclusive thinking across research, education, administration and leadership.

A central aim is to provide a common platform for exchange. By bringing together researchers, administrators and decision-makers, the Group encourages dialogue across roles and disciplines. This collaborative setting supports the development of organisational cultures and research practices in which inclusion is treated not as an add-on, but as a guiding principle. The ambition is to move from short-term initiatives towards structural change that becomes part of everyday practice.

The creation of the Working Group also strengthens ERCIM's wider mission. It signals a clear commitment to responsible research and innovation, enhances institutional reputation and supports international competitiveness. By increasing the visibility of existing work and enabling the sharing of best practice, the Group helps address persistent imbalances in participation across ICT disciplines. It also aligns closely with European Union policy priorities and funding frameworks that link excellence in research with social impact.

Strategically, the Group contributes to the development of human capital within ERCIM. It supports the Fellowship Programme by helping to attract and retain talented researchers from a wide range of backgrounds. It also positions ERCIM as a European reference point for inclusive excellence in digital research. Through channels such as ERCIM News, the Group's activities can reach a broad audience, extending their influence well beyond the immediate network and strengthening connections among member organisations.

Research

Research is a core pillar of the Working Group's activity. Members are encouraged to develop collaborative projects, organise workshops and events, and pursue interdisciplinary work that addresses both technical and cultural dimensions of inclusion. Indicative areas include human computer interaction, digital accessibility, end user applications for the web, mobile platforms and gaming, digital health, cyber security, social media and discrimination, artificial intelligence, machine learning, ambient intelligence, the internet of things and big data analysis and management, with particular attention to intersectional perspectives.

Education and awareness raising form a second pillar. The Group aims to support early career researchers, facilitate the exchange of expertise across ERCIM members and communicate research insights to the wider public. In addition, it seeks to contribute to policy development by producing roadmaps and guidelines that help organisations and research teams make their practices and outcomes more inclusive.

By linking technology and culture, the Inclusive Digital Futures Working Group highlights that progress in digital systems depends as much on organisational and societal change as on technical innovation. Its long-term vision is to ensure that future technologies advance scientific and economic goals while contributing to fairer and more representative digital societies.

Invitation to join

Researchers, practitioners, administrators and leaders from across ERCIM and beyond who are interested in shaping inclusive digital futures are warmly invited to join the Working Group and contribute to its activities. For further information or to express interest, please contact the Working Group Chair.

Link:
https://ercim-ide-wg.ics.forth.gr/

Please contact:
Magdalini Chatzaki
FORTH ICS, Greece
Working Group chair
magda@ics.forth.gr

ERCIM Forum "Beyond Compliance" 2025

by Anaëlle Martin (CCNEN – French National Digital Ethics Advisory Committee)

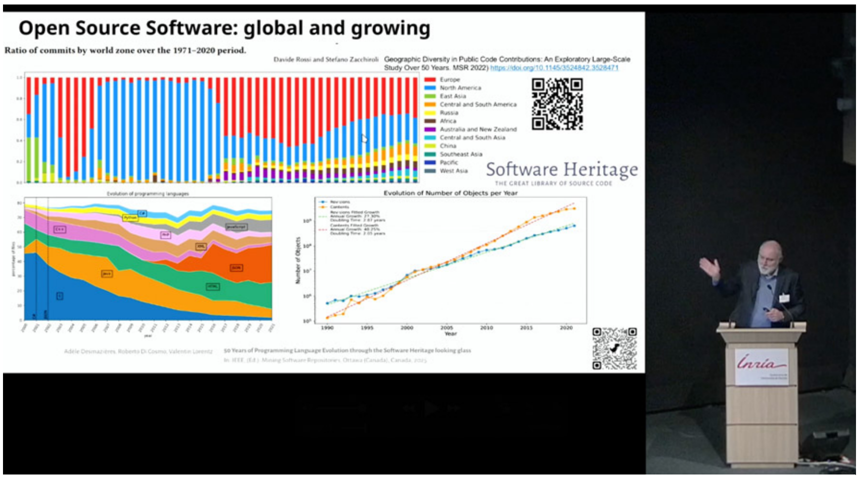
The fourth edition of the ERCIM Forum "Beyond Compliance" [L1] was held from 29 to 31 October 2025 at the Inria Centre in Rennes. Returning to France after the inaugural edition in Paris in 2022 [L2], the Forum once again provided an international platform for discussion on ethical issues in digital research and innovation.

The three-day programme addressed five key themes: security in the digital society; geopolitics of digital ethics; data altruism and open academic resources; generative AI in research, teaching and publishing; and AI's impact on behaviour and cognition. Hosted by the Inria Centre at Rennes University and IRISA, the 2025 edition reaffirmed ERCIM's commitment to debate, critical reflection and interdisciplinary collaboration.

The three-day programme addressed five key themes: security in the digital society; geopolitics of digital ethics; data altruism and open academic resources; generative AI in research, teaching and publishing; and AI's impact on behaviour and cognition. Hosted by the Inria Centre at Rennes University and IRISA, the 2025 edition reaffirmed ERCIM's commitment to debate, critical reflection and interdisciplinary collaboration.

Wednesday, 29 October

The Forum opened with a joint round table with the Ethics Working Group of Informatics Europe, held in connection with the Informatics Europe ECSS event in Rennes [L3]. Presentations addressed the responsible development and application of information systems, with particular attention to accountability, misinformation, ethical awareness and best practices in digital security. Discussions focused on compliance and accountability in AI development, the impact of design choices on user auton-



Roberto Di Cosmo from Inria and University Paris Cité, presenting "No science without source: collecting, preserving and sharing software in a risky world" during the session "Data altruism and open academic resources" on Thursday, 30 October 2025.

omy, and ethical risks related to generative AI, misinformation and deepfakes.

The day concluded with a keynote by Catherine Tessier on ethical issues in AI research and education. The keynote examined how the rapid diffusion of generative AI challenges fundamental notions of learning, writing, responsibility and scientific practice, and called for a rethinking of research and educational practices rather than simple adaptation to AI tools.

Thursday, 30 October

The day began with a strategy meeting of the ERCIM Digital Ethics Working Group, followed by a tutorial by Alexei Grinbaum on AI's impact on human behaviour and cognition, drawing on results from the AIOLIA project [L4]. The tutorial highlighted both potential cognitive benefits and risks such as automation bias and overreliance, stressing the importance of operationalising ethical principles through actionable, context-sensitive guidelines.

The late morning session focused on data altruism and open academic resources. Contributions examined free and open-source software as a form of digital altruism, large-scale preservation infrastructures such as Software Heritage [L5], and the role of legal and governance constraints, including GDPR [L6]. The discussion also highlighted the European Open Science Cloud (EOSC) [L7] and its relationship to FAIR data practices [L8]. Speakers emphasised that openness alone does not guarantee trustworthiness and that responsible data sharing requires contextualisation, transparency and human-centred design. The session also referenced recent recognition of work in this area [L9].In the afternoon, a keynote by Afonso Seixas-Nunes addressed the moral and legal responsibilities associated with autonomous systems, particularly autonomous weapons, highlighting challenges related to accountability, proportionality and the opacity of AI-based systems.The day concluded with a session on the geopolitics of digital ethics in academia, analysing how power, sovereignty and international tensions shape ethical practices in universities and research institutions. Speakers referred to major international and European policy developments, including the Council of Europe's framework convention [L10], the EU AI Act [L11], and relevant UN-level discussions [L12].

Friday, 31 October

The final day focused on the integration of generative AI in academic practices. Speakers critically examined the largely unreflective adoption of generative AI in research and education, highlighting pedagogical, scientific, social and environmental costs, and questioning which intellectual tasks should remain protected from automation.The Forum concluded with a keynote by Mihalis Kritikos on ethical considerations in EU-funded digital projects under Horizon Europe [L13]. The keynote presented the Commission's approach to embedding ethics as a learning-oriented, ongoing process rather than a purely compliance-driven exercise, while warning against fragmentation and "ethics washing".

Outlook

Videos of the Forum talks will be made available online soon. Readers interested in being notified when the recordings are published, and in future Beyond Compliance activities, are invited to subscribe to the Beyond Compliance mailing list [L14].

Links:
[L1] https://www.ercim.eu/beyond-compliance
[L2] https://www.ercim.eu/beyond-compliance/beyond-compliance-2022
[L3] https://www.informatics-europe.org/ecss/2025/home
[L4] https://aiolia.eu/index.php/partners/
[L5] https://www.softwareheritage.org/author/roberto/
[L6] https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng
[L7] https://kwz.me/hDX
[L8] https://www.eoscobservatory.eu/explore/open-science-by-area/fair-data
[L9] https://kwz.me/hGX
[L10] https://kwz.me/hGW
[L11] https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng
[L12] https://docs.un.org/en/A/79/l.118
[L13] https://kwz.me/hGZ
[L14] https://lists.ercim.eu/wws/subscribe/beyondcompliance

Please contact:
Anaëlle Martin
CCNEN – French National Digital Ethics Advisory Committee, France
anaelle.martin@ccne.fr

First Workshop of the ERCIM Working Group on Quantum Technologies

by Carlo Mastroianni (CNR-ICAR) and José Francisco Chicano García (UMA)

On 2 December 2025, the inaugural workshop of the newly established ERCIM Working Group (WG) on Quantum Technologies (QT) was held at the Lloyd's Baia Hotel in Vietri sul Mare, Italy. Organised as a central part of the XXI ICAR-CNR workshop, this event brought together experts and stakeholders to discuss the future of quantum computing, communication, and sensing. The meeting was sponsored by the National Quantum Science and Technology Institute (NQSTI) and supported by the newly formed WG, marking a significant milestone in ERCIM's commitment to the emerging quantum economy.

Fabio Martinelli, Director of CNR-ICAR, addresses the participants at the ERCIM event on Quantum Computing, held as part of the XXI ICAR–CNR Workshop on 2 December 2025.

Strategic Vision of the Working Group

The establishment of the QT Working Group is a response to the rapid reshaping of information processing and sensing driven by quantum technologies. The working group, currently chaired by Francisco Chicano (University of Malaga) and Carlo Mastroianni (CNR), aims to strengthen European research by bringing together expertise from academia and industry to address critical bottlenecks such as scalability, error correction, system integration and benchmarking.

Guided by its chairpersons and a Steering Committee (SC), which is still under development, the Group aims to accelerate the transition from laboratory-based experiments to deployable quantum technologies. This synergy is essential for Europe's leadership in the emerging quantum economy. More specifically, the Working Group aims to focus on a series of activities such as projects, workshops and knowledge dissemination for fostering the European research and development on quantum computing, communication, sensing and hardware platforms. These will be among the main issues of current and future research efforts for "quantum technologies" in a broad sense in Europe [L1, L2]). One main source of possible funding will be EU research programmes. In particular, the scope of the WG matches the objectives of the Quantum Flagship [L3] and the funding opportunities offered by EU calls [L4].

Morning Session: Future Trends in Academia

The workshop [L5] opened with introductions from Emilio Fortunato Campana (Director of the DIIET Department of CNR) and Han La Poutré (CWI and Elected President of ERCIM). The morning sessions focused on research trends within academia and research institutes, featuring contributions from a wide range of ERCIM and external institutions including NQSTI (Italy, Francesco Saverio Cataliotti), University of Malaga (Spain, Francisco Chicano), FORTH/University of Crete (Greece, George Stamatiou), King's College of London (UK, Mohammad Reza Mousavi), CNR (Italy, Carlo Mastroianni), NTNU (Norway, Asle Sudbø), SBA (Austria, Sebastian Raubitzek), and RISE (Sweden, Miroslav Dobsicek).

Key research areas discussed included:
- Foundations and Algorithms: investigating quantum algorithms for optimisation, information retrieval, and machine learning.

The workshop organised by ICAR-CNR was held in Vietri sul Mare, Italy on the beautiful Amalfi coast.

- Networking and Security: exploring Quantum Key Distribution (QKD) and the protocols necessary for a future Quantum Internet.
- Hybrid Systems: analyzing the interplay between quantum hardware and classical High Performance Computing (HPC).
- Quantum Software Engineering: adapting and creating new methodologies to design, test, and debug quantum and hybrid systems.

The academic dialogue was further enriched by a contribution from the European Commission, represented by Prof. Oscar Diez (DG CONNECT) who highlighted the alignment between the WG's goals and major European initiatives such as the Quantum Flagship and the development of the European Quantum Communication Infrastructure (EuroQCI).

Afternoon Session: Industrial Contributions

The afternoon session, chaired by Marco Pota (ICAR-CNR) was dedicated to industrial research innovation trends. It featured an in-depth exchange with representatives from major industrial players and specialised quantum technology companies, including D-Wave, Quantum2PI, Leonardo, IQM, Fujitsu, QuantumNet and QTI.

The goal of this session was to identify industrial challenges that can be tackled through academic research and vice versa. Industry experts and researchers discussed the development of diverse hardware platforms, such as superconducting, photonic, and ion-trap systems, while emphasizing the need for best practices in scientific reproduction and performance benchmarking.

Key Outcomes and Future Initiatives

The workshop was characterised by intense engagement and productive networking, including several collaborative initiatives on specific scientific themes, which will help to define the WG's trajectory in the coming years. A major focus of the discussions concerned joint participation European research funding programmes. The WG is actively encouraging its members to prepare joint proposals targeting opportunities under the Quantum Flagship and other EU funding schemes. Furthermore, there was a clear consensus on extending the group's membership. Plans are already underway to expand the WG beyond its initial core to include other ERCIM institutions, external research entities, and industrial partners. This expansion is intended to create a more comprehensive and inclusive ecosystem for quantum research in Europe.

In addition to these research goals, dissemination also represents a key mission. A Special Issue of ERCIM News dedicated to Quantum Technologies is currently under preparation, aiming to showcase recent scientific advances and strategic developments within the community. Building on the success of this inaugural event, planning has already begun for the next workshop, which is expected to further consolidate the group's role as a cornerstone of European quantum innovation.

Conclusion

The first workshop of the Quantum Technology Working Group successfully established a platform for innovation and collaboration. By supporting the growth of early career researchers, stimulating mobility, and strengthening links between academia and industry, the WG is well positioned to drive the transition of quantum technologies from theory to reality. The foundations laid in Vietri sul Mare promise a vibrant and productive future for the ERCIM quantum community.

Researchers interested in contributing to the activities of the Working Group are warmly invited to participate and are encouraged to contact the chairs for further information.

Links:
[L1] https://qt.eu/
[L2] https://kwz.me/h8I
[L3] https://qt.eu/about-quantum-flagship/
[L4] https://qt.eu/funding-opportunities/
[L5] https://ercim-qt-wg.icar.cnr.it/

Please contact:
Carlo Mastroianni, CNR-ICAR, Italy
Working Group co-chair
carlo.mastroianni@icar.cnr.it

Francisco Chicano
University of Málaga, Spain
Working Group co-chair
chicano@uma.es

A New ERCIM Working Group on Artificial Intelligence & Intelligent Systems (AIIS)

by Giuseppe Manco (CNR) and the AIIS Working Group founding members

Artificial Intelligence (AI) is rapidly reshaping the way we build software, run critical infrastructures, and deliver public services. From healthcare and mobility to cybersecurity and environmental monitoring, AI systems are becoming core components of Europe's digital ecosystem. At the same time, the growing adoption of data-driven and foundation-model technologies is exposing major scientific and societal challenges: ensuring robustness, transparency, fairness, accountability, and alignment with European values and regulations.

To address these challenges in a coordinated, transnational manner, ERCIM has launched the Working Group on Artificial Intelligence and Intelligent Systems (AIIS). The AIIS WG is conceived as a collaborative forum where ERCIM members, together with academic, industrial, and public stakeholders, can jointly advance both the fundamentals and the real-world impact of AI research.

Why an ERCIM AIIS Working Group now?

AI innovation is accelerating, but progress is not only a matter of performance improvements. Increasingly, AI research is required to address questions such as: Can we trust the behaviour of complex models? Can we understand their decisions? Can they be evaluated rigorously across contexts and populations? Can they meet EU regulatory requirements while remaining useful and scalable?

The AIIS WG was created precisely to strengthen ERCIM's collective capacity to respond to these questions. Its ambition is to help grow a visible and active ERCIM AI community, to provide a structured channel for joint workshops and thematic seminars, to stimulate competitive European project proposals, and to enable mobility exchanges and shared research initiatives, supported by a public-facing presence and regular reporting.

In line with the ERCIM spirit of the Working Groups, the AIIS WG is designed to build and maintain a network of researchers in the AI and intelligent systems domain. Participation is open to anyone who wishes to contribute, while ensuring a stable ERCIM core (at least three ERCIM member organisations).

Purpose and scope

The AIIS WG focuses on fundamental and applied AI, with a strong emphasis on trustworthy, explainable, and human-centric intelligent systems.

From a methodological standpoint, the WG embraces a broad set of paradigms, including:
- Machine learning and deep learning
- Reinforcement learning and optimisation
- Symbolic AI and knowledge representation
- Hybrid neuro-symbolic systems and cognitive architectures.

A distinctive element of the WG is its explicit commitment to the development of trustworthy AI, including fairness, transparency, robustness, explainability, and ethical alignment, as well as scalable assessment of normative compliance with EU regulations.

The scope also includes emerging directions that are central to the next generation of AI systems, such as multimodal learning, data-efficient and continual learning, and the evaluation, training, and application of Large Language Models (LLMs) and Visual Language Models (VLMs).

On the systems side, AIIS addresses the integration of AI with enabling technologies such as IoT, edge computing, robotics, digital twins, scalable AI infrastructures, and energy-efficient deployment approaches for resource-constrained settings.

Finally, the WG explicitly welcomes AI research that is validated through meaningful application domains, including cybersecurity, smart cities, telecommunications, health, bioinformatics, mobility, and environmental challenges.

Activities and expected outcomes

In line with ERCIM Working Group principles, AIIS will be structured around four interconnected activity axes: workshops, projects and proposals, mobility, and dissemination/visibility.

A cornerstone activity is the organisation of at least one international workshop per year, open to researchers in the field. These events will combine scientific exchange with agenda-setting, including identifying research challenges, consolidating collaborations, and shaping joint initiatives that can evolve into publications and project proposals.

The WG will also actively foster joint participation in European funding programmes such as Horizon Europe and Digital Europe, and will contribute to building strategic partnerships that connect ERCIM expertise with the needs of industry and public bodies.

Mobility and talent development represent another key pillar. AIIS will encourage short research visits across member institutes, shared PhD topic development, and contributions to the ERCIM Fellowship Programme through attractive, high-impact research topics.

A growing community: current participants and partner mapping

The AIIS WG currently brings together an initial nucleus of ERCIM institutes including CNR (Italy), FORTH (Greece), HUN-REN SZTAKI (Hungary), Inria (France), and ITIS-UMA (Spain).

The AIIS Working Group warmly welcomes expressions of interest from ERCIM institutes and the wider research community. As AI becomes increasingly central to Europe's competitiveness and resilience, AIIS aims to provide an ERCIM platform where excellent research, practical impact, and trustworthy innovation can advance together.

Please contact:
Giuseppe Manco, CNR-ICAR, Italy
giuseppe.manco@icar.cnr.it

ERCIM “Alain Bensoussan” Fellowship Programme

The ERCIM Postdoctoral Fellowship Programme is one of the flagship initiatives of ERCIM. Open to young researchers from around the world, the programme covers a broad range of fields in computer science and applied mathematics.

> “The ERCIM Fellowship has been a pivotal milestone in my academic career. It provided an exceptional platform for interdisciplinary research, international collaboration, and professional development. Through this fellowship, I had the opportunity to engage with leading experts across Europe, expand my research horizons, and gain valuable insights that have significantly influenced my long-term academic and professional objectives. The experience has not only enriched my scholarly perspective but also strengthened my global research network.
>
> **Pallab Kumar NATH**
> Former ERCIM Fellow

The fellowship scheme aims to help young scientists deepen their knowledge of European research structures and networks, while gaining valuable experience within leading European research institutions. Fellowships have a duration of 12 months, with the possibility of extension, and are hosted by one of the ERCIM member institutes.

Hosting Institutions

Only ERCIM members can host fellows. When an ERCIM member is a consortium, the hosting institute may be any of its member organisations. When an ERCIM member is a funding body, the hosting institute may be one of its affiliated institutions.

Fellowships are offered according to the needs and available funding of the member institutes.Fellows are appointed either through a stipend (a research training agreement) or a work contract, depending on the hosting institute. The type of contract and the monthly allowance or salary vary by host institution.

ERCIM encourages applications from researchers both in academia and in industry.

Why apply for an ERCIM Fellowship?

The Fellowship Programme enables talented early-career scientists from all over the world to work on challenging research problems alongside leading European experts. In addition to research excellence, the programme fosters collaboration and knowledge exchange within the European research community.

The programme offers ERCIM Fellows the opportunity to:
- Work with internationally recognised experts;
- Gain a deeper understanding of European research structures and networks;
- Become familiar with the working conditions in leading European research centres;
- Promote cross-fertilisation and cooperation between research groups working in similar areas across Europe.

Equal Opportunities

ERCIM is committed to ensuring equal opportunities and promoting diversity. Candidates are not discriminated against on the basis of race, colour, religion, gender, national origin, age, marital status, or disability.

Conditions

Candidates must:
- Have obtained a PhD degree within the last eight years (prior to the application deadline), or be in the final year of their doctoral studies with an outstanding academic record. Proof of the PhD degree must be provided before the start of the fellowship;
- Be fluent in English.

Application deadlines

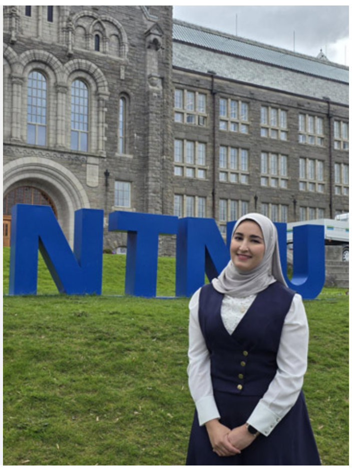
Applications are accepted twice a year, with deadlines on 30 April and 30 September.

Since its inception in 1991, more than 800 fellows have participated in the programme.

The Fellowship Programme is named in honour of Alain Bensoussan, former president of Inria, one of the three founding institutes of ERCIM.

http://fellowship.ercim.eu

> “The fellowship has boosted my professional profile, expanded my international research network, and provided new opportunities for collaboration. It has helped me refine my technical skills, strengthen my publication record, and position me for future grant opportunities—all of which will be invaluable in advancing my academic and research career.
>
> **Roufaida LAIDI**
> Former ERCIM Fellow



Introduction to the Special Theme

AI for Science

Advancing Discovery, Enhancing Trustworthiness, and Reshaping Scientific Practices

by the guest editors Edina Nemeth (SZTAKI) and Alexandre Termier (University of Rennes – Inria/IRISA, France)

Artificial Intelligence (AI) is rapidly transforming the way science is conducted. From accelerating the discovery of new materials to modeling complex climate systems and supporting biomedical research, AI has become an essential tool for advancing knowledge. By enabling more efficient data analysis, powerful simulations, and new forms of hypothesis generation, AI helps researchers tackle problems that were previously too complex or time-consuming to solve.

In line with the European vision of trustworthy, human-centred, and sustainable AI, this special theme explores the growing role of AI as a catalyst for scientific progress across disciplines. It aims to showcase innovative methods, interdisciplinary collaborations, and real-world applications where AI enhances responsible scientific research and discovery. Contributions highlight both fundamental advances in AI technologies and their application to pressing scientific challenges.

AI as a catalyst for the research process

The articles selected in this section focus on AI systems that directly support the two ends of the research process: the selection of research topics and the evaluation of research through peer review.

In their article *Enhancing Reviewer Idenfication with AI*, Berndt et al. present an AI-based reviewer identification system that automates the matching of manuscripts to suitable experts by leveraging semantic analysis of abstracts and publication histories. The approach helps ease bottlenecks in peer review and improving the quality, efficiency and fairness of scientific evaluation (page 14).

From a different perspective, the article *AI Assistant for Research Topic Selection in Higher Education*, Stampfl et al. illustrates how large language models can be embedded in higher education as Socratic tutors, guiding students and early-career researchers from vague ideas to consistent, methodologically sound research proposals while demanding conceptual clarity and coherence (page 16).

Together, these contributions show AI being used to strengthen the research cycle end-to-end, from early topic design to the assurance of quality in publication.

Building trustworthy AI for scientific decision support

With AI now driving high-stakes decisions, trustworthiness is no longer optional — it's a core scientific priority.

McAleer et al., in *Developing Human-Centred Trustworthy AI as Infrastructure for Reliable Decision Support*, report on THEMIS 5.0, which designed methods for assessing accuracy, robustness, and fairness with domain experts in healthcare, maritime operations, and journalism (page 17).

Beecks et al., in *A Human-Centred AI Approach to Data-Driven Scientific Discovery*, present a human-centred data science framework based on Gaussian process models which enable the extraction of interpretable, uncertainty-aware insights while keeping the data scientist in control of the discovery process (page 19).

The contribution by Sapidis et al. introduces *SemanticRAG: Traceable Answers from Documents and Knowledge Graphs*, which combines retrieval-augmented generation with knowledge graphs and explicit provenance, enabling question answering in which every claim is traceable back to specific document snippets or graph triples (page 20).

In the article *Foundation Models and Trustworthy AI for Environmental Systems*, Hatzivasilis et al. examine foundation models through the lens of security assurance and continuous risk assessment, showing how trustworthy AI and security-by-design can be integrated into large-scale monitoring and control platforms (page 22).

Across these contributions, trustworthiness is treated as a measurable, context-dependent property that must be engineered, monitored, and enhanced.

Re-engineering the building blocks of AI

A series of articles examine the computational substrates and resource consumption of AI itself.

Biological Reservoir Computing experiments described by Ciampi et al., demonstrate that living neuronal cultures, interfaced through high-density multi-electrode arrays, can act as physical reservoirs for pattern recognition, pointing toward bio-hybrid neuromorphic architectures that combine rich dynamics with potential gains in energy efficiency and adaptability (page 24).

Barbierato et al., in *Blockchain Energy Costs for AI-Driven Scientific Infrastructure*, quantify the energy costs and trade-offs of blockchain-based infrastructures that are often proposed as trustworthy backbones for AI-driven scientific platforms, highlighting that verifiability and decentralisation have a tangible environmental price (page 25).

Complementing this in a second contribution, Barbieto et al. argue that in scientific contexts *Accuracy is not Enough*, and that computational efficiency and alignment with domain knowledge remain key design objectives when deploying AI at scale to ensure reproducibility and value in gaining additional insights (page 27).

Together, these works emphasise that advancing AI for science also requires rethinking its basic components to make them more sustainable, explainable, and physically grounded.

AI for imaging, sensing, and scientific observation

Many scientific advances depend on extracting meaning from visual and spatial data, an area where AI is rapidly becoming indispensable.

Szentirmai et al. discuss *Lightweight on Device AI* showing how resource-aware models running on augmented reality devices can act as flexible scientific infrastructure, supporting universally designed experiments and studies in the field (page 28).

Computational Imaging research presented by van Leeuwen et al. explores how AI can push beyond traditional image reconstruction pipelines, enabling new modalities, reducing computational cost, and improving image quality in resource-constrained settings (page 30).

The article presented by Ćeranić et al. describes a hybrid method for *Detecting Small Changes in 3D Aerial Scans*, including structural or terrain changes of real urban environments (page 31).

Burnet and Parisot report on *Fully Automated Detection of Harmful Cyanobacteria Blooms in Lakes Using Photo Traps and YOLO-based Object Detection*, enabling near-real-time early warning in critical water reservoirs (page 33).

These contributions exemplify how AI expands our observational capabilities, turning heterogeneous, high-volume image streams into operational scientific insight.

AI for society and for modelling complex systems

Finally, the special theme highlights AI's role in modelling complex systems and supporting societal missions.

An *AI-enhanced Operational Picture for Public Safety Operations* combines heterogeneous textual sources with knowledge graphs to build semantically coherent situational overviews for emergency responders, addressing both information overload and domain heterogeneity is presented by Pöschl et al. (page 34).

Segura Ortiz et al. show how *Context-guided Evolutionary Algorithms* are transforming gene regulation inference by integrating multiple computational approaches with biological knowledge (page 35).

In energy and climate-related modelling, a generative adversarial surrogate trained on high-fidelity simulations learns to reproduce wind-turbine wakes at hub height, offering orders-of-magnitude speed-ups that enable optimisation and control studies that would otherwise be computationally prohibitive (Toutouh et al, page 37).

Work on deep learning-enhanced multi-scale simulations of molecular systems shows how AI can bridge modelling scales, coupling atomistic detail with mesoscopic or continuum descriptions in ways that preserve essential physics while reducing computational burden (Christofi and Hamandaris, page 39).

Across these efforts, AI serves as a mediator between data and theory, enabling models that are fast enough, and expressive enough, to be embedded into real-world decision processes.

Concluding remarks

Taken together, the articles in this special theme make clear that AI technologies are no longer peripheral tools in scientific research. AI for Science represents an evolving ecosystem of tools, infrastructures, and practices that cut across disciplines. They demonstrate AI helping to organise the scientific process, deepening trust in model-based decisions, facilitating the understanding of complex systems, extending our senses, and enabling new forms of modelling and control. At the same time, this transformation raises fundamental questions about trust, transparency, energy consumption, and the evolving role of human expertise in the scientific process.

We hope this collection will not only inform readers about current advances, but also inspire new collaborations at the interface of AI and the many scientific disciplines it now helps to transform.

As we witness the rapid development of AI technologies themselves, we are also witnessing their growth into a fundamental scientific infrastructure underpinning discovery across disciplines. Emergent approaches such as Denario or Sakana [L1] even aim to automate key steps in scientific discovery — from hypothesis generation to experimental design. This suggests that we may be only at the beginning of a deeper transformation in how scientific knowledge is generated.

Link:
[1] https://github.com/AstroPilot-AI/Denario

Please contact:
Edina Nemeth, SZTAKI, Hungary
edina.nemeth@sztaki.hun-ren.hu

Alexandre Termier
University of Rennes – Inria/IRISA, France
alexandre.termier@irisa.fr

Enhancing Reviewer Identification with AI

by René Berndt, Hillary Farmer and Eva Eggeling (Fraunhofer Austria)

Improving the reviewer selection process for conferences and journals using AI and large language models (LLMs) can significantly enhance both efficiency and quality. AI-driven systems can analyse manuscripts and match them with potential reviewers based on their expertise, publication history, and prior reviewing experience. By leveraging semantic understanding rather than relying solely on manually assigned keywords, LLMs enable more accurate and nuanced reviewer–paper alignment.

The Importance of Publishing in Scientific Research

In scientific disciplines, and particularly within research contexts, the publication of results constitutes an essential aspect of scholarly activity. Sharing research findings with other researchers not only advances knowledge but also contributes to the scientific community's collective understanding. The reputation of a researcher is largely determined by bibliometric data, such as the venues where their work is published (impact factor) and the frequency with which it is cited by other researchers like the h-Index (Hirsch index or Hirsch number).

One key process for ensuring the quality of scientific papers is peer review. In this process, a submitted manuscript is evaluated by other experts from the same field, commonly referred to as peers. Peer review is a fundamental component of scientific publishing, as it guarantees the rigor, validity, and credibility of scientific work. By relying on experts to assess research, the process upholds high standards and enhances the quality of published studies.

Challenges in the Peer Review Process

Effective peer review is dependent on the careful matching of manuscripts with qualified reviewers. This task is critical but time-consuming. With the continual increase in the number of scientific publications, finding suitable reviewers and managing the peer review process has become cumbersome and demands significant effort.

The task of finding the most suitable reviewer has been addressed in systems like the Toronto Paper Matching System (TPMS) [2], which relies on reviewers uploading their own publications. This works fine if you have a closed pool of possible reviewers, usually the programme committee of a conference. But when you need additional reviewers, which are not part of that pool, these systems cannot provide this information. SARA – Services for Aiding Reviewer Assignment suggests reviewers using only manuscript abstracts. Unlike TPMS, SARA leverages LLM embeddings to compare abstract semantics with those in its database, without referencing full publication texts.

Data backbone

The prototype system uses the monthly datasets from DBLP[L1], a large and openly accessible bibliographic database that focuses on computer science research. However, DBLP does not include abstracts. This is where the Initiative for Open Abstracts (I4OA) becomes relevant. I4OA represents a partnership among publishers, infrastructure organizations, librarians, researchers, and others who work together to promote free and open access to the abstracts of scholarly works—especially journal articles and book chapters—in trusted repositories that are both open and accessible to machines and submit them to Crossref [1].

The data backbone of SARA contains over 2 million publications with abstract with 1.8 million authors (see Fig. 1).

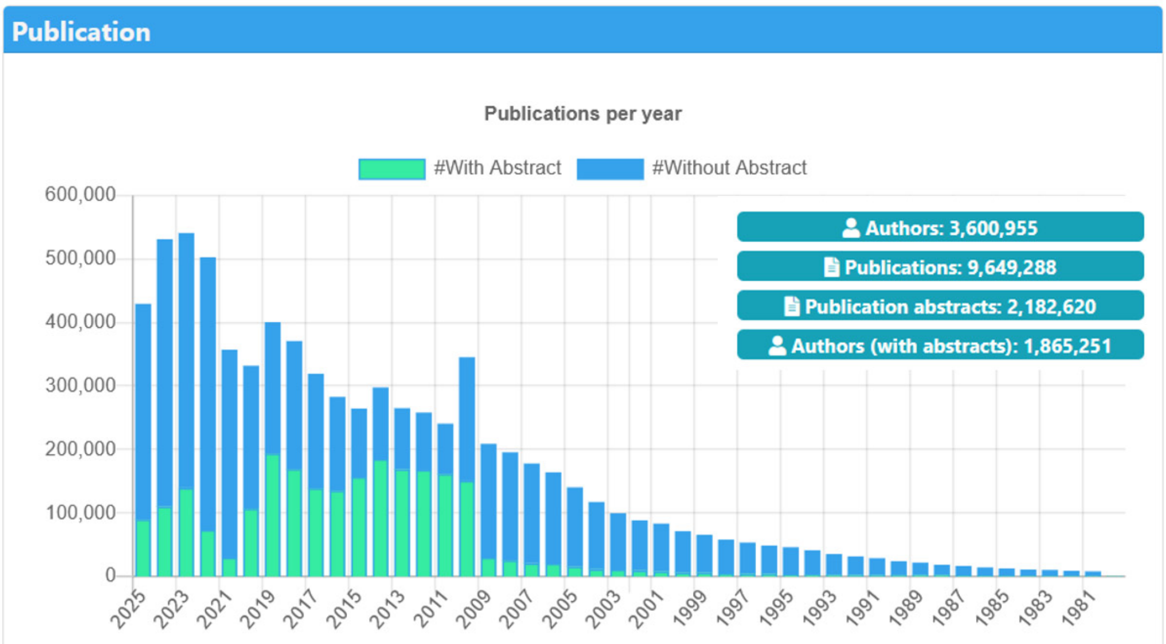


Figure 1: Overview of the SARA data backbone based on DBLP and I4OA, showing the distribution of publications with available abstracts versus records without abstracts.

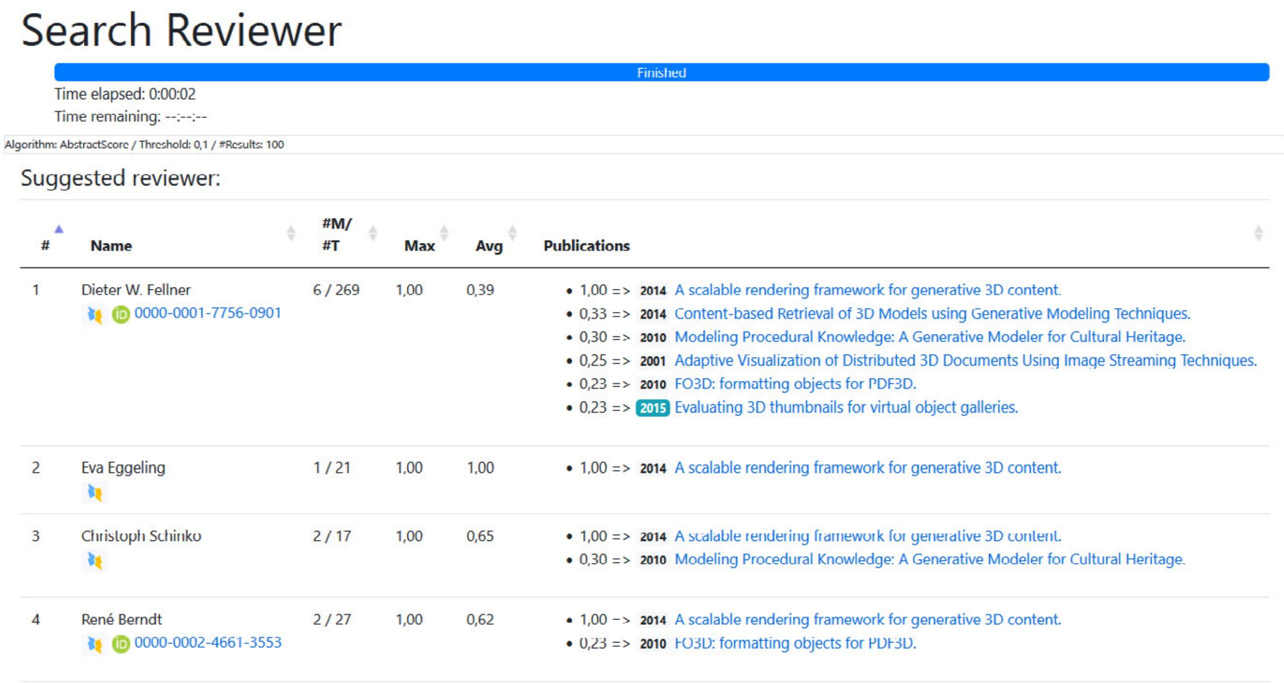


Figure 2: Example output of the SARA reviewer identification system integrated into SRMv2, presenting ranked reviewer candidates with similarity metrics, publication statistics, and links to ORCID profiles and DOI-referenced publications.

Integration with SRMv2

SARA has been integrated in SRMv2 [L2], the submission and reviews system of the Eurographics – a professional association founded in 1980 that promotes research, development, and education in computer graphics across Europe and internationally.

The functionality is accessible via a single link, which sends the abstract to SARA. The resulting output is a list of individuals qualified to serve as reviewers for the submission. Various metrics are provided to assist in identifying the most suitable candidates according to different criteria (see Figure 2).

It provides several details, including the name, DBLP link, ORCID page, and additional information:
- #M – Number of matching publications of the person above the threshold
- #T – Total number of publications
- Max: The maximum similarity score of the matching publications
- Avg: The Average of the similarity scores of the various publications.

The publication column shows matched titles and years by similarity score, with each title linked to its official repository, typically via DOI.

This information assists editors or conference chairs in selecting appropriate reviewers, as these metrics indicate the reviewer's level of expertise within a particular field (number of relevant publications) and their overall seniority (total number of publications). For example, a candidate who closely matches the abstract but has published only one or two papers may be less suitable than another with over thirty publications, including ten directly related to the topic of the abstract.

Links:
[L1] https://dblp.org/
[L2] https://srmv2.eg.org

References:

[1] I4OA: Initiative for Open Abstracts, [Online]. Available: https://i4oa.org/. Accessed: Oct. 12, 2025.
[2] L. Charlin and R. Zemel, "The Toronto Paper Matching System: An automated paper-reviewer assignment system," May 2013. [Online]. Available: https://www.cs.toronto.edu/~lcharlin/papers/tpms.pdf. Accessed: Nov. 19, 2025.

Please contact:

René Berndt, Fraunhofer Austria
rene.berndt@fraunhofer.at

AI Assistant for Research Topic Selection in Higher Education

by Rita Stampfl (University of Applied Sciences Burgenland), Barbara Geyer (University of Applied Sciences Burgenland)

At the University of Applied Sciences Burgenland, a GPT-based chatbot has been developed to support students in creating research topic proposals for scientific papers. Large Language Models like GPT-4 enable interactive conversations, allowing chatbots to facilitate complex learning processes and provide personalised learning experiences. In the rapidly changing educational landscape, specially designed educational chatbots are gaining importance. This trend, combined with the accessibility of Large Language Models and the ability to create GPTs without programming knowledge, opens new possibilities for integrating them into learning environments. To ensure these chatbots function as intelligent tutors rather than simple question-answer machines, appropriate instruction is essential.

To understand the context, it is worth looking briefly at historical developments. Chatbots have been present in the education sector since the 1960s and have evolved from simple response devices to highly sophisticated learning assistants. Kuhail et al. [1] describe this evolution from teaching agents to interactive ones. According to them, the majority of chatbots have used chatbot-driven conversation to guide users through pre-structured dialogues, whereas only a minority have relied on user-driven conversation in which learners control the interaction through AI-supported responses. LLMs like ChatGPT have fundamentally changed the capabilities of educational chatbots and made user-controlled conversations possible. Thanks to these models, chatbots can now conduct significantly more complex and demanding conversations than was possible with pre-programmed responses. These so-called Socratic chatbots use targeted questions to promote critical thinking and self-reflection instead of providing direct answers. This creates a dialogue-oriented learning approach that expands learners' abilities to support complex learning processes and enables personal learning experiences.

When developing a GPT as a research topic assistant [L1], the learning objectives were first identified and the needs of the target group analysed. Based on this, a didactic concept was developed that integrates the creation of a scientific paper as the core of the learning interaction. The focus is on questioning strategies that promote critical thinking and autonomous learning. The GPT was designed as a role-play game to leverage the advantages of game-based learning. These enable learning as a combination of theory and application in authentic situations [2]. Role-play games demonstrably increase student interest in learning. The GPT prompt was newly developed following the methodology described by Stampfl and Prodinger [3], who demonstrated how ChatGPT can be effectively employed as an assistant for topic selection in scientific papers.

The "Themendispo Assistant" operates according to a structured dialogue principle. Students are systematically guided through various dimensions of their research project. The interaction begins with identifying the overarching topic area and systematically delves into more specific aspects such as problem formulation, resulting research gaps, goal definition, and scientific questioning. The theoretical background, choice of methods, and anticipated results are also critically examined. Through targeted questions, the tutor encourages users to consider and challenge the structure of their research project. This not only promotes understanding of the topic but also trains analytical thinking and problem-solving skills.

A particular strength of the "Themendispo Assistant" lies in its ability to evaluate logical consistency between elements of the research design. It accepts only precise, scientifically grounded answers. In cases of inconsistency or insufficient precision, it requests revision. The communicative style of the Themendispo Assistant is deliberately formal and demanding. At the same time, the constructive support function is not neglected. This requires a high degree of scientific rigour and conceptual clarity from students, corresponding to academic standards of excellence. The assistant supports learners in expanding their knowledge through dynamically generated ques-

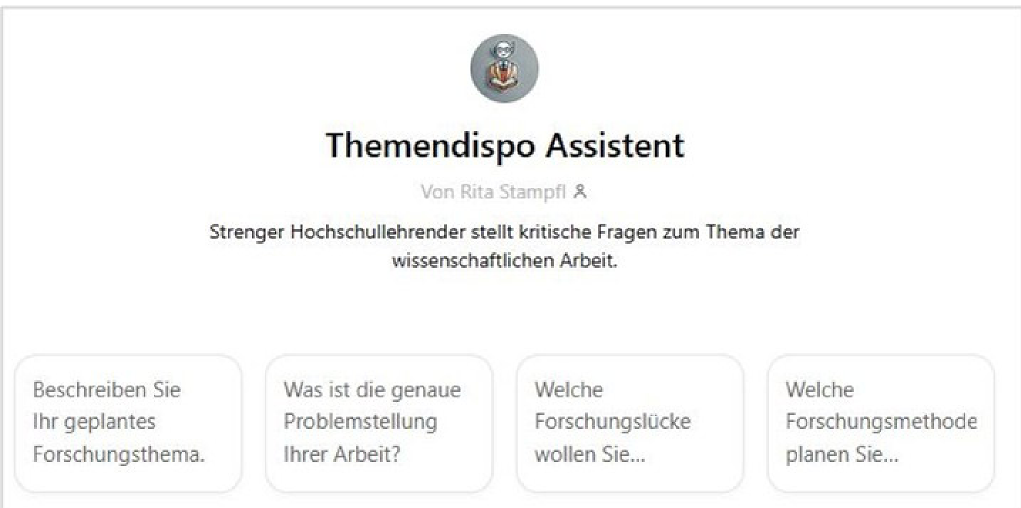


Figure 1: Collaborative XR learning environment as a basis.

tions that build on users' previous answers, creating a personalised learning experience that encourages independence.

The prototypes of these tutors were created using the GPTs in GPT-4 and evaluated in various test phases. Initial testing was conducted in a controlled environment to verify the technical and didactic effectiveness. User feedback was collected to measure the quality of interaction and learning success. This feedback was crucial for iterative adjustments to both the target group approach and functional design of the chatbot. Ongoing tests and adjustments served to increase the effectiveness of tutors and ensure a personalised learning experience that actively supports users in their learning processes.

The GPT was designed as a complement to conventional supervision. The Themendispo Assistant represents progress in applying artificial intelligence in education and illustrates the potential of adaptive dialogue systems for quality assurance in academic training. Through systematic guidance in developing consistent research concepts, the system promotes scientific competence among students while reducing demands on academic staff. The application of Socratic methods in chatbots stimulates a dialogue between teachers and learners that goes beyond the mere retrieval of information and supports the development of a sound understanding of the subject matter. Future research could examine the long-term effects of such AI tutors on the quality of scientific work.

Links:
[L1] https://kwz.me/hIC
[L2] https://openai.com/blog/introducing-gpts
[L3] https://barbarageyer.substack.com
[L4] https://www.linkedin.com/in/barbara-geyer/

References:

[1] M. A. Kuhail et al., "Interacting with educational chatbots: A systematic review," Education and Information Technologies, vol. 28, no. 1, pp. 973-1018, 2023. https://doi.org/10.1007/s10639-022-11177-3
[2] J. Matute Vallejo and I. Melero, "Learning through play: The use of business simulators in higher education teaching," Universia Business Review, vol. 51, pp. 72-111, 2016. https://doi.org/10.3232/UBR.2016.V13.N3.03
[3] R. Stampfl and M. Prodinger, "KI-Planspiel zur Themendisposition: ChatGPT als Assistent zur Themenfindung fur wissenschaftliche Arbeiten," R&E-SOURCE, vol. 11, no. 4, pp. 119-130, 2024. https://doi.org/10.53349/resource.2024.i4.a1345

Please contact:

Rita Stampfl
University of Applied Sciences Burgenland, Austria
rita.stampfl@hochschule-burgenland.at

Developing Human-Centred Trustworthy AI as Infrastructure for Reliable Decision Support

by Susie Ruston McAleer (21c) and Spiros Borotis (Maggioli S.p.A)

THEMIS 5.0 is generating evidence, tools, and methods that help scientific communities evaluate when AI systems can be trusted. Through pilots in healthcare, maritime operations, and journalism, the project is exploring how trustworthiness assessments can help organizations take up AI in a responsible manner.

As AI becomes central to modelling, prediction, and analysis, work in domains of critical societal importance increasingly relies on systems that operate as collaborators rather than mere computational tools. But such reliance requires that AI models are used responsibly. Hence, organizations need to know: How accurate is the model? How stable is it under different conditions? Is it fair across groups? What risks might it amplify? Without rigorous answers, AI can introduce hidden errors into critical work processes, distort outputs, or produce misleading guidance.

THEMIS 5.0 [L1] addresses this emerging challenge. Instead of treating trustworthiness as an abstract principle, the project investigates how scientists and practitioners understand trustworthiness and what it takes to allow for decisions to be based on model outputs. Towards this end, the project has involved users and stakeholders within healthcare, port operations, and journalism in requirements identification, scenario development, and feedback on trustworthiness components. These activities produced empirical insights, which illuminates what people truly need to use AI responsibly and confidently in their research and work. Figure 1 illustrates how technical performance and user priorities together shape trust in AI-supported decision making [1].

Protecting the Integrity of Decision Making with AI

Reliable performance is the primary concern for professionals using AI. In THEMIS 5.0 user and stakeholder involvement, clinicians have stressed that high global accuracy is essential not only for patient safety but also to ensure that downstream research and decisions are based on stable predictions. Journalists highlighted how accuracy varies across languages, noting that cross-lingual inconsistencies can distort content analysis and misinformation research. Port operators similarly emphasised robustness, the ability of models to remain stable under changing conditions, as a prerequisite for reproducible simulation and planning.

In response, research in THEMIS 5.0 integrates tools to assess trustworthiness characteristics, at the level of individual samples. This allows for a tiered approach to accuracy assessment, augmenting assessments of global and group-specific accuracy with assessments of accuracy for individual samples.

Preventing Bias from Entering Decisions

Fairness is a critical requirement for trustworthy AI, as systematic differences in model behaviour across groups can distort data, misrepresent populations, and propagate bias into scientific research. In the feedback from users and stakeholders gathered through THEMIS 5.0, clinicians stressed the need to distinguish genuine medical variation from algorithmic bias, while port operators warned that preferential treatment of certain vessels or companies could undermine economic and logistics research. Journalists similarly noted that fairness failures in media AI can skew studies of online behaviour and misinformation.

To address these risks, THEMIS 5.0 is developing a Fairness Evaluator to provide clear, contextual explanations rather than isolated statistical metrics from fairness assessments. This narrative approach helps users and stakeholders identify whether observed disparities arise from data, modelling choices, or contextual factors, supporting more accurate and representative analysis.

Making the Use of AI Responsible and Transparent

The requirements for AI decision support are shaped by users and stakeholders' roles, responsibilities, and values. THEMIS 5.0 has explored this through the concept of a Persona Analyser, which identifies what users and stakeholders of different domains prioritise when assessing AI trustworthiness. User and stakeholder involvement in the project shows that for preference modelling to support trustworthy AI, it must be accessible and non-intrusive. Users and stakeholders require clear, high-level summaries of their preferences with an option to examine details on demand.

Drawing on these insights the Persona Analyser communicates preferences in straightforward, qualitative terms, for example, that a researcher prioritises fairness over robustness, helping align trustworthiness evaluation with scientific responsibilities without burdening users. This ensures that AI evaluation remains transparent and appropriate to the context in which decisions are made.

Making AI Failures Visible

Many AI-related risks are not visible through accuracy or fairness metrics alone. In THEMIS 5.0, the risk modelling tool Spyderisk [L2] helps users and stakeholders see how risks arise from the interaction of models, data, and real environments. The tool is useful for the governance of the trustworthy AI assessment process: it may in particular be essential for pre-deployment checks and for validation of deployed AI decision making support, but may also have potential use for operational application. THEMIS is now refining Spyderisk with clearer narrative explanations so teams can conduct risk assessment and mitigate threats before they influence evidence or decisions.

A Clearer Picture of Trustworthy AI

Across all domains, THEMIS shows that AI used for decision support must be subject to trustworthiness assessment not only for safe deployment, but also to safeguard the reliability of knowledge produced through AI-supported work. Inaccurate models can distort data, unfair models can misrepresent populations, fragile models undermine reproducibility, and unseen risks can compromise both decisions and the scientific insights derived from them. By providing interpretable trustworthiness assessments, THEMIS enables users and stakeholders to judge when AI outputs are reliable enough to inform decisions and analyses in practice. In this way, the project contributes to the development of human-centred trustworthy AI as a foundational capability for scientific and evidence-informed decision making, without compromising accuracy, fairness, or public trust. [2]

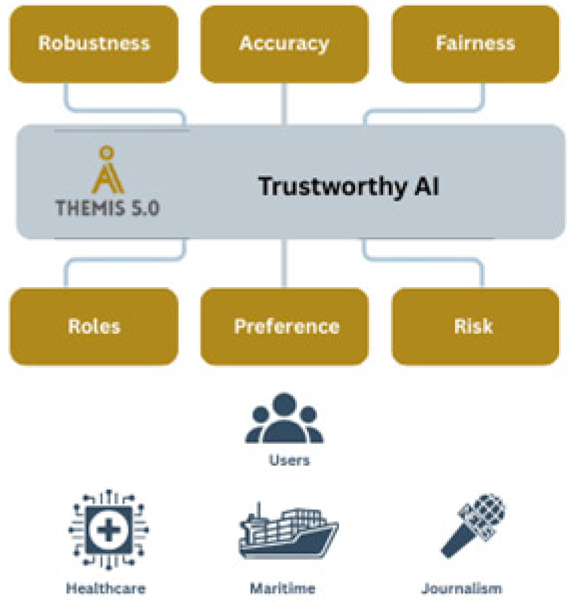


Figure 1: Elements of Trustworthy AI.

THEMIS 5.0 has received funding from the EU Horizon Europe Research and Innovation programme under grant agreement No. 101121042, and from UKRIs funding guarantee.

Links:
[L1] www.themis-trust.eu
[L2] www.it-innovation.soton.ac.uk/projects/spyderisk

References:
[1] P. V. Johnsen, et al., "SPARDACUS SafetyCage: A new misclassification detector," in Proc. 6th Northern Lights Deep Learning Conf. (NLDL), Proc. Mach. Learn. Res., vol. 265, pp. 133–140, 2025. [Online]. Available: https://proceedings.mlr.press/v265/johnsen25a.html
[2] M. Fikardos, et al., "Trustworthiness optimisation process: A methodology for assessing and enhancing trust in AI systems," Electronics, vol. 14, no. 7, Art. no. 1454, 2025, doi: 10.3390/electronics14071454.

Please contact:
Susie Ruston McAleer
21c, United Kingdom
susie@21cConsultancy.com

A Human-Centred AI Approach to Data-Driven Scientific Discovery

by Christian Beecks (FernUniversität in Hagen) and Markus Lange-Hegermann (Technische Hochschule Ostwestfalen-Lippe)

Data-driven scientific discovery increasingly relies on artificial intelligence. This article presents a human-centred data science framework based on Gaussian process models which enable the extraction of interpretable, uncertainty-aware insights while keeping the data scientist in control of the discovery process.

Data science is a modern scientific discipline that researches data-analytical methods for complex problem solving. Being positioned at the frontiers across math, computer science, and artificial intelligence, data science has evolved into an interdisciplinary methodology for data-driven scientific discovery. This new era of scientific discovery, which is also known as the fourth paradigm of science [1], is characterized by big versatile data landscapes and advanced analysis methods from the fields of machine learning and artificial intelligence. Revealing actionable and interpretable insights from big data and complex algorithms is hence one of the major challenges in data science.

Apart from these challenges, the interdisciplinary nature of data-driven scientific discovery necessitates an efficient means of implementation. While a broad range of data technologies and analysis methods are domain-agnostic, the scientist's aims and abilities are crucial when designing a data-driven solution to a specific data scientific problem. Prior knowledge and domain expertise are a prerequisite when narrowing down the data analytical solution space to apt machine learning models and artificial intelligence methods.

The process of data-driven scientific discovery typically requires extensive manual effort. Following the conceptual analysis workflow of the data science life cycle [2], data first undergoes cleaning, preparation and exploratory analysis before statistical inference and model estimation are adopted to domain-specific research questions. Though automated machine learning approaches have mitigated the time-consuming task of building, tuning and deploying machine learning models, these approaches follow a strictly formalised objective function, making it difficult to uncover insights hidden in the data.

The extraction of interpretable and, where appropriate, actionable insights is one of the greatest data science capabilities, particularly when applied in an open-ended manner, where the research question or hypotheses need to be first developed and then empirically verified. In this case of scientific discovery, structured hypotheses based on data effects and hidden phenomena, such as trends and correlations as well as dependencies and causalities need to be determined efficiently based on limited human effort. To mitigate information complexity during this process, raw data is typically abstracted by means of data models. Among the broad variety of data models, such as symbolic regression, simple causal models and interpretable regression models, Gaussian process models have turned into a general-purpose model widely used in Bayesian machine learning. Not only their stochastic nature, but also their ability to dynamically cope with incomplete and uncertain data in an interpretable way, shows that Gaussian processes are well suited for trustworthy data-driven scientific discoveries.

Our aim is to exploit Gaussian process models for the development of a trustworthy and interpretable approach to data-driven scientific discovery. While recent methods towards data science automation make extensive use of large language models [3], we aim to exploit a clear Bayesian formalism by means of Gaussian process to (i) model data in an optimal way, (ii) derive transparent analytical insights, and (iii) deliver comprehensible explanations. These three pillars define a new human-centred perspective on data science that epitomizes a trustworthy and interpretable approach for data-driven scientific discovery.

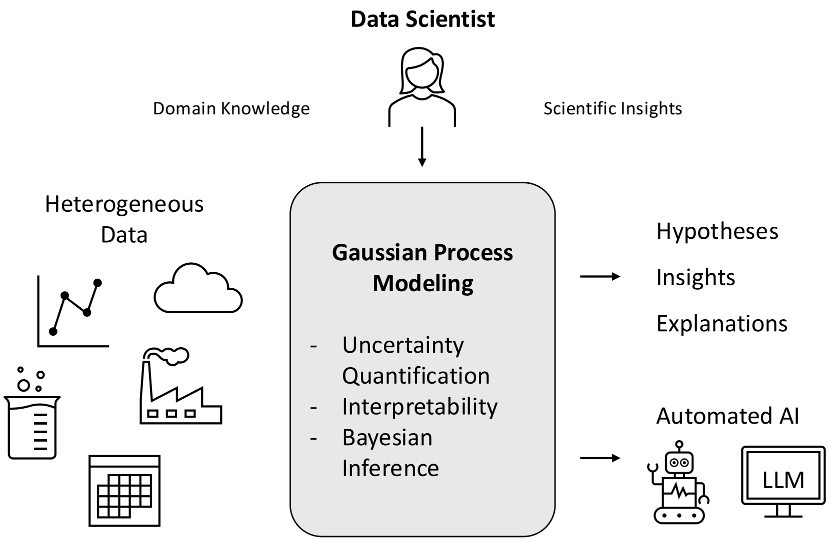


Figure 1: Human-centred framework for trustworthy AI in data-driven scientific discovery. Gaussian process models form an interpretable and uncertainty-aware core that connects heterogeneous data sources with scientific insights, while allowing human experts to guide hypothesis generation, model refinement, and interpretation. Automated AI components support analytical tasks without replacing human control.

Coming back to its origin, the last decade has advanced data science from a code-centric approach, where data scientists act as programmers implementing analytical operations line-by-line, to a no-to-low-code approach, where data scientists leverage automation frameworks powered by automated machine learning to perform analytical operations. Based on scientific innovations in the field of generative artificial intelligence, data science has evolved into an artificial intelligence-driven approach [3], where agentic systems orchestrated by large language models try to guess the data scientist's intention and execute analytical operations. While the latter approaches show remarkable performance in typical data science tasks ranging from data preparation to report generation, they are also able to conduct deep research on diverse data sources [3].

Against this background, we propose a methodological framework that makes use of Gaussian process models as a core component for trustworthy data-driven scientific discovery. Instead of replacing the data scientist with fully automated agentic systems, our approach augments human expertise by providing transparent probabilistic models that support hypothesis generation, validation, and exploitation. By explicitly quantifying uncertainty and enabling interpretable model structures, Gaussian processes allow data scientists to explore complex data landscapes while maintaining control over analytical decisions. This combination of Bayesian inference and human-centred artificial intelligence fosters reproducible insights and supports responsible scientific discovery across diverse application domains.

References:

[1] S. Tansley and K. M. Tolle, The Fourth Paradigm: Data-Intensive Scientific Discovery, T. Hey, Ed. Redmond, WA, USA: Microsoft Research, 2009.

[2] V. Stodden, "The data science life cycle: A disciplined approach to advancing data science as a science," Commun. ACM, vol. 63, no. 7, pp. 58–66, Jul. 2020.

[3] S. Zhang et al., "DeepAnalyze: Agentic large language models for autonomous data science," arXiv preprint, arXiv:2510.16872, 2025.

Please contact:

Christian Beecks
FernUniversität in Hagen, Germany
christian.beecks@fernuni-hagen.de

Markus Lange-Hegermann
Technische Hochschule Ostwestfalen-Lippe, Germany
markus.lange-hegermann@th-owl.de

SemanticRAG: Traceable Answers from Documents and Knowledge Graphs

by Iordanis Sapidis, Michalis Mountantonakis and Yannis Tzitzikas (FORTH-ICS and University of Crete)

SemanticRAG [1] is an interactive QA system that answers questions using both documents and Knowledge Graphs. To mitigate the black-box nature of LLMs, it provides provenance for every answer, citing the exact document snippet or KG triple from which it originates so users can verify each claim.

Access to large document corpora, such as those found in digital libraries, is typically supported through services including browsing, keyword search, and faceted search. Recently, Retrieval-Augmented Generation (RAG) methods have been introduced, leveraging Large Language Models (LLMs) to enable Question Answering (QA) while mitigating the hallucination problem in such models. Building on this development, this paper investigates an approach for enabling QA over document corpora and Knowledge Graphs (KGs) by integrating LLMs, RAG, and RAG enhanced with Knowledge Graph information. To address the challenge of black-box interaction, we present SemanticRAG, an interactive system that allows users to pose questions, compare answers generated by different methods, and examine the provenance of each response.

An overview of the online demo is shown in Figure 1, where SemanticRAG receives the question, "Is the freshwater sculpin protected by the EU Habitats Directive?". The system produces three answers: one from an LLM, one from RAG and one from RAG enhanced with data from KGs (KG RAG). In this setup, RAG can access a corpus of documents on ecosystem restoration, while KG RAG can access the same documents plus a fish taxonomy. We observe that the LLM asserts that the freshwater sculpin is not a protected species, and the RAG approach reports insufficient information to answer. By contrast, KG RAG returns the correct answer by retrieving (i) a document stating that the species Cottus gobio is protected and (ii) a row in the verbalized KG indicating that "freshwater sculpin" is another name for Cottus gobio. Consequently, KG-RAG can answer questions not only about the documents or the KG in isolation, but also those that require synthesizing information from both sources.

To accomplish this, our system uses KGVerbalizer, a tool that converts KGs into natural language sentences using rule-based templates defined in a JSON configuration. This enables us to verbalize heterogeneous knowledge resources, such as taxonomies and ontologies, and index those sentences alongside document passages in a single corpus. At query time, a unified retriever searches this corpus and retrieves the most relevant snippets, whether they come from a KG or from a document, within one pipeline. We combine lexical similarity, which captures exact terms and entities, with semantic similarity, which captures paraphrases and contextual matches, and then apply a reranking step to keep only the most relevant candidates. The resulting top passages are passed to the answer generator.

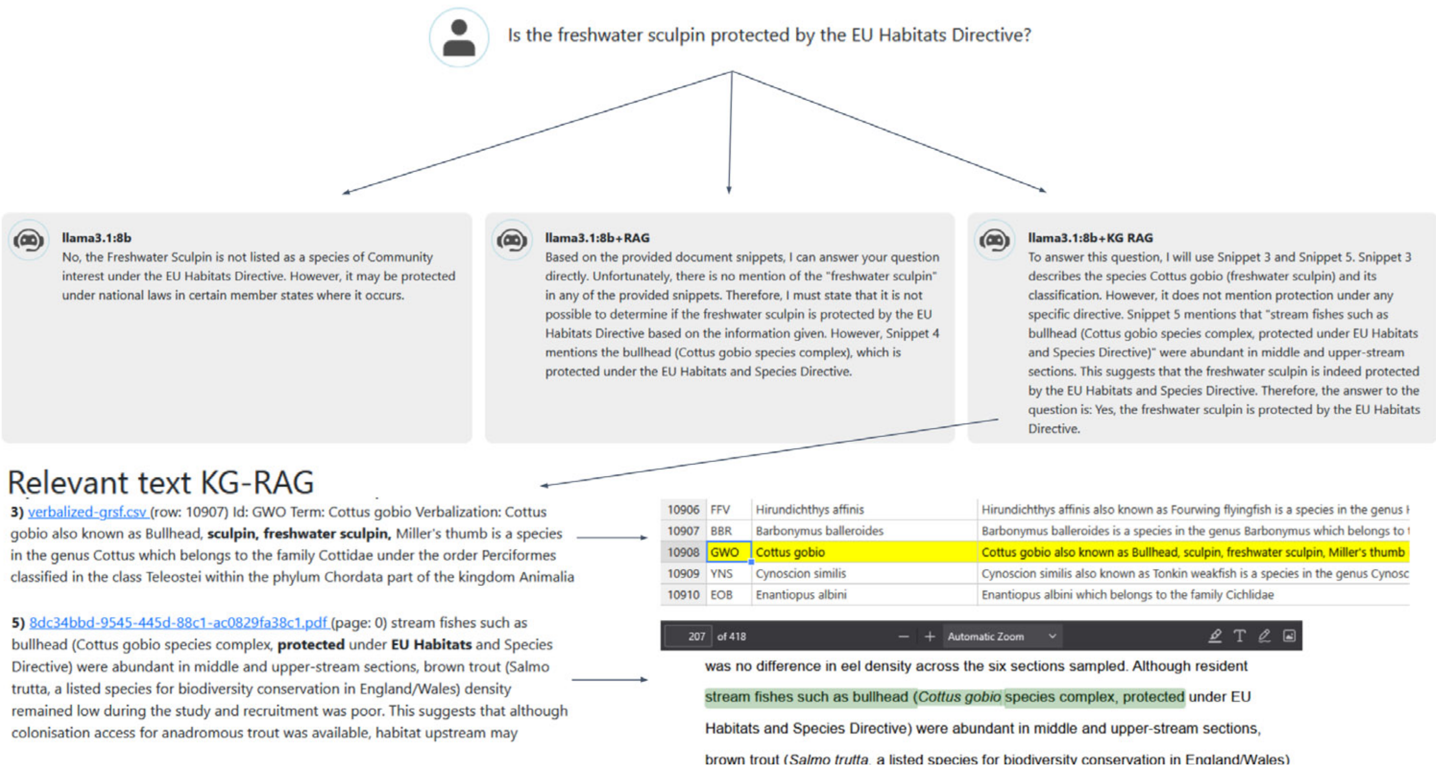


Figure 1: Running example of SemanticRAG.

The SemanticRAG online demo [L1] is deployed on a corpus comprising papers on ecosystem restoration collected by the FAO (Food and Agriculture Organization) and a GRSF (Global Record of Stocks and Fisheries) fish taxonomy. The interface exposes two configurable parameters, top-k and score_threshold: the former specifies the number of snippets retrieved and supplied to the LLM for response generation, while the latter filters snippets by relevance.

Regarding evaluation, we assessed the SemanticRAG pipeline using both QA benchmarks and a task-based user study. On the CRAG benchmark [L2], our system outperforms a standard RAG baseline, achieving 18.6% higher accuracy and a 22.1% reduction in hallucination rate. For KG RAG, we evaluated on MetaQA [2] and WC2014QA [3]; the model attains satisfactory performance on MetaQA (Hits@1 = 88.9%) and near-perfect performance on WC2014QA (99.3%) without any training. In the user study, KG-RAG surpassed both the standalone LLM and the vanilla RAG system in answer correctness and user satisfaction by at least 20% on each metric.

As future work, we plan to (i) evaluate the impact of different verbalization rules at each stage on QA performance; (ii) assess KG-RAG on benchmarks that integrate both unstructured documents and KGs; and (iii) develop methods to extract and index information from images and tables for incorporation into the retrieval process.

Links:
[L1] https://demos.isl.ics.forth.gr/SemanticRAG

References:

[1] I. Sapidis et al., "Interactive and provenance-aware search and QA over documents using LLMs, RAG and knowledge graph verbalization," in Proc. TPDL, 2025.
[2] L. Zhang et al., "Gaussian Attention Model and Its Application to Knowledge Base Embedding and Question Answering," arXiv:1611.02266, 2016
[3] Y. Zhang et al., "Variational Reasoning for Question Answering with Knowledge Graph," in Proc. AAAI, 2018.

Please contact:

Iordanis Sapidis, FORTH-ICS and University of Crete
sapidis@ics.forth.gr

Michalis Mountantonakis
FORTH-ICS and University of Crete
mountant@ics.forth.gr

Yannis Tzitzikas, FORTH-ICS and University of Crete
tzitzik@ics.forth.gr

Foundation Models and Trustworthy AI for Environmental Systems

by George Hatzivasilis and Sotirios Ioannidis (Technical University of Crete) and François Hamon (Greencityzen)

Environmental sciences increasingly rely on AI foundation models to integrate heterogeneous data and support sustainable decision-making. Their real-world impact, however, depends on security, trustworthiness, and resilient deployment, as illustrated by secure smart watering systems.

Environmental Sciences in the Era of Foundation Models

Environmental systems are among the most complex domains addressed by contemporary computing. Phenomena such as water availability, soil conditions, air quality, and climate dynamics emerge from interactions across multiple spatial and temporal scales and are observed through highly heterogeneous data sources. These include in-situ sensor networks, satellite imagery, meteorological forecasts, and numerical simulations.

Traditional machine learning approaches, typically trained for a single task or dataset, struggle to cope with this complexity. In response, AI foundation models have emerged as a transformative paradigm. Trained on large, diverse datasets, foundation models learn general representations that can be adapted to multiple downstream tasks. In environmental sciences, this enables unified modelling of physical processes, contextual reasoning across data modalities, and efficient adaptation to new regions or conditions.

For applications such as water management (see Figure 1), foundation models enable irrigation strategies that go beyond fixed rules. Instead, decisions can be informed by learned relationships between soil moisture, weather evolution, seasonal patterns, and historical behaviour, improving both sustainability and resilience under increasing climatic variability.

Yet, while foundation models offer unprecedented analytical power, their deployment outside controlled research environments introduces new risks.

From Environmental Intelligence to Operational Systems

Environmental AI increasingly influences cyber-physical systems. Smart irrigation platforms connect sensors, AI-driven decision logic, and automated actuators in continuous feedback loops. In these settings, AI outputs directly trigger physical actions, such as opening or closing irrigation valves.

This transition exposes a critical gap. Many AI systems are evaluated primarily for predictive accuracy, with limited consideration of how they behave under adversarial conditions, component failures, or malicious interference. In operational environments, such omissions can lead to resource waste, service disruption, or safety risks.

To be viable in practice, environmental AI systems must therefore satisfy three fundamental properties:
- Security: to resist cyber threats across devices and software
- Trustworthiness: to ensure explainable and reliable decisions
- Resilience: to maintain functionality despite faults or attacks.

These requirements are especially pressing for systems that rely on foundation models, which are often reused, updated, and integrated across multiple services.

Smart Watering as a Representative Use Case

Smart watering systems [L1] exemplify this convergence of environmental intelligence and operational risk. Distributed sensors monitor soil and environmental conditions, AI-based logic determines irrigation needs, and actuators execute watering actions automatically. When operating correctly, such systems reduce water consumption and improve plant health.

However, these infrastructures are also exposed. Compromised sensors may report misleading data, manipulated controllers can trigger excessive irrigation, and unauthorised access to AI components can distort decision-mak-

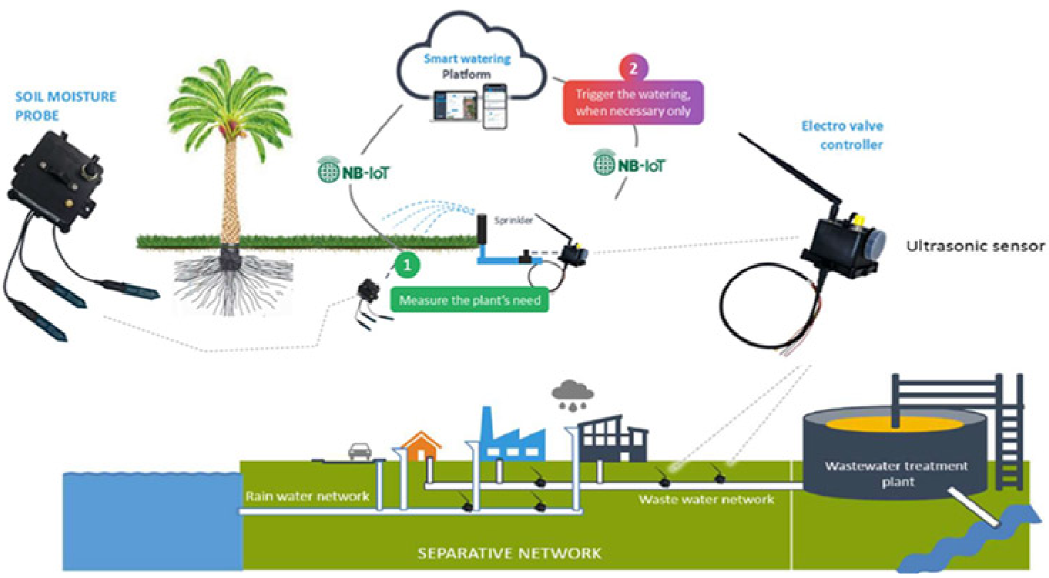


Figure 1: Water management system.

ing in subtle but impactful ways. Addressing these risks requires systematic security assessment that explicitly accounts for AI-driven functionality.

Layer-Aware Security Assessment with SecOPERA

The SecOPERA framework [1] addresses this challenge by providing a layer-aware security assurance process that decomposes a system into device, network, application, and cognitive layers, enabling targeted assessment of each component (see Figure 2). This approach is particularly significant for AI-enabled environmental systems. Rather than treating machine learning models as opaque artefacts, SecOPERA explicitly recognises them as cognitive assets that must be assessed and protected.

The cognitive security assessment focuses on the integrity and protection of AI decision-making at inference time. Since many AI components are implemented as opaque deep neural networks, the assessment does not depend on exposing internal model parameters. Instead, it evaluates the inference-layer exposure of the AI system (e.g., serving APIs, edge deployments, or automated decision pipelines), analysing risks related to unauthorised access, model extraction/inversion attempts, adversarial manipulation, and abnormal query behaviour that could undermine reliable decision-making.

These analyses are integrated with broader security testing across the system, ensuring that AI-driven decisions are evaluated in the context of the physical processes they control.

From Assessment to Trustworthy Deployment

A defining strength of the SecOPERA approach is its closed-loop assurance workflow. Assessment results are collected in structured reports and evaluated against predefined security objectives [2]. When weaknesses are identified, developers can apply targeted hardening measures, such as reducing unnecessary dependencies, integrating security-assured modules, or reconfiguring exposed interfaces.

Crucially, AI components are re-assessed after adaptation, supporting iterative improvement rather than one-off validation. This process enables continuous monitoring and response across distributed environmental infrastructures and supports long-term operational trust.

Conclusion

Foundation models are becoming indispensable to environmental sciences, enabling integrated reasoning across complex, multi-scale systems. Their societal value, however, depends on more than analytical performance. Without embedded security and trustworthiness, even the most advanced environmental AI cannot be responsibly deployed.

The smart watering use case demonstrates how foundation-model–enabled intelligence, combined with layer-aware security assessment and cognitive-layer analysis, can support sustainable and reliable environmental services. By treating AI models as security-relevant components rather than black boxes, frameworks such as SecOPERA provide a practical path toward trustworthy environmental AI.

As environmental decision-making increasingly relies on automated intelligence, security-by-design will be essential, not as an afterthought, but as a foundation for responsible innovation.

Link:
[L1] https://youtu.be/jZrJWNAhx4M

References:

[1] A. Fournaris et al., "Providing Security Assurance & Hardening for Open Source Software/Hardware: The SecOPERA approach", IEEE CAMAD 2023.
[2] M. Papoutsakis et al., "SESAME: Automated Security Assessment of Robots and Modern Multi-Robot Systems", Electronics, MDPI, vol. 14, issue 5, article 923, pp. 1-27.

Please contact:

George Hatzivasilis
Technical University of Crete (TUC), Greece
gchatzivasilis@tuc.gr

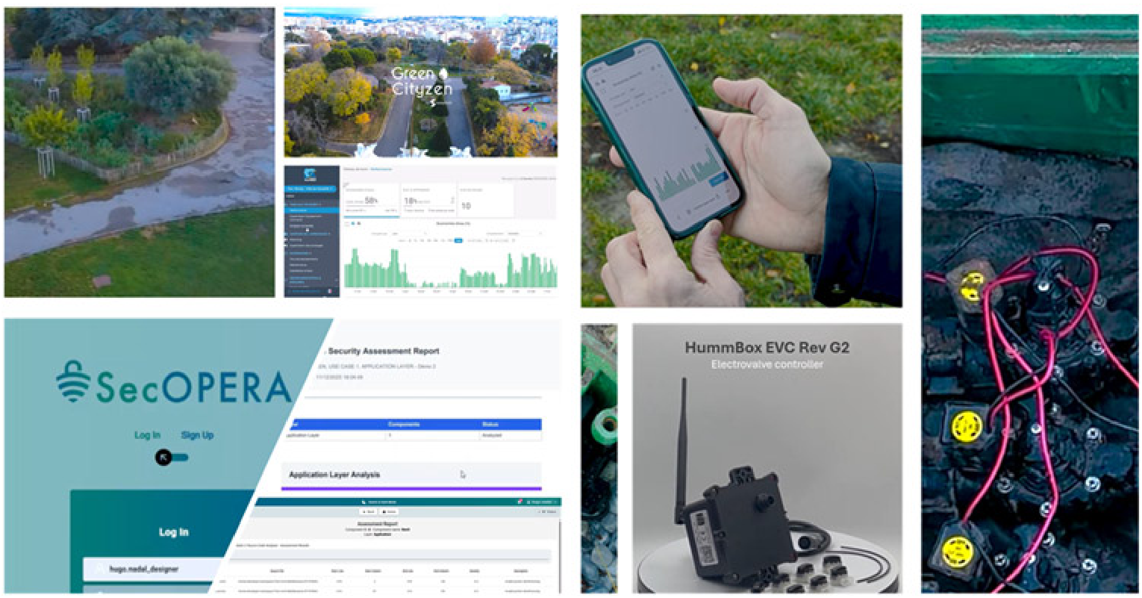


Figure 2: SecOPERA use case of smart water management.

Biological Reservoir Computing: Harnessing Living Neurons for AI

by Luca Ciampi, Ludovico Iannello, Giuseppe Amato (CNR-ISTI), Federico Cremisi and Fabrizio Tonelli (Scuola Normale Superiore Pisa)

Can living neurons compute? Researchers from CNR-ISTI, CNR-IBF, and Bio@SNS introduce a pioneering approach in which cultured neuronal networks act as reservoirs for pattern recognition. This bio-hybrid paradigm aims to bridge neuroscience and machine learning, opening new pathways towards interpretable and energy-efficient AI.

Artificial Intelligence (AI) has achieved remarkable progress through deep learning, yet most current models remain abstract approximations of biological neural systems. This gap has inspired growing interest in biologically grounded approaches that combine machine learning principles with the intrinsic dynamics of living neurons. Our work explores this frontier by introducing Biological Reservoir Computing (BRC)—a paradigm in which a network of cultured neurons serves as the computational substrate for AI tasks. The project brings together computer scientists and biologists from CNR - ISTI, CNR-IBF, and Bio@SNS in Pisa, Italy, combining expertise in machine learning, electrophysiology, and stem cell biology to explore this innovative direction. These activities have been carried out within the framework of the NRRP project THE: Tuscany Health Ecosystem.

What is Biological Reservoir Computing?

Reservoir Computing (RC) is a machine learning framework that projects input data into a high-dimensional space through the nonlinear dynamics of a recurrent system. This transformation makes patterns easier to separate, so that even a simple linear classifier can achieve good performance without training the reservoir itself. Traditionally, RC implementations rely on artificial units, such as Echo State Networks or Liquid State Machines. In contrast, BRC replaces these artificial reservoirs with real neuronal networks cultured in vitro, leveraging their natural complexity and nonlinear dynamics to process information.

How does it work?

The system interfaces with a high-density multi-electrode array (HD-MEA), which enables both electrical stimulation and high-resolution recording of neural activity. Input patterns—such as digit-like configurations—are encoded as spatial stimulation sequences and delivered to the neuronal culture. The evoked spiking responses are captured across thousands of electrodes and transformed into high-dimensional feature vectors. These representations are then classified using a simple linear readout layer. This approach effectively turns a living neural network into a biologically instantiated feature extractor. Unlike conventional artificial reservoirs, the biological substrate operates with intrinsic variability and rich dynamics, offering a unique perspective on neuromorphic computation. The overall framework is shown in Figure 1.

Why is this important?

Modern AI systems face challenges related to energy consumption, interpretability, and biological plausibility. By offloading part of the computation to a physical neural substrate, BRC offers potential advantages in energy efficiency and adaptive behavior, while providing insights into how real neural networks process information. This line of research also contributes to the broader effort of bridging neuroscience and AI, in line with the European vision of trustworthy and sustainable AI.

Experimental Results

To validate the feasibility of BRC, we conducted experiments on pattern recognition tasks of increasing complexity. Starting from simple geometric patterns, we progressed to clock-digit-like configurations and finally to real handwritten digits from the MNIST dataset. Despite the inherent variability of biological responses, the system consistently produced discriminative representations that enabled accurate classification. In our

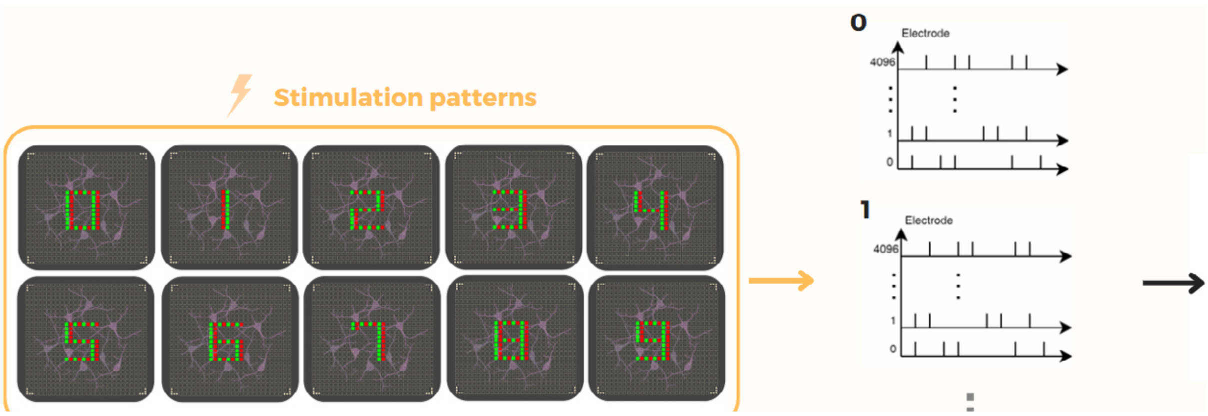


Figure 1: Overview of the Biological Reservoir Computing (BRC) concept. A multi-electrode array (MEA) acts as a two-way interface to a cultured biological neural network, allowing us to both stimulate the living neurons and record their responses. Inputs are represented by activating specific electrodes, which deliver controlled electrical pulses to the network. The resulting neural activity is captured through other electrodes and converted into rich, high-dimensional patterns that encode the input in a latent computational space. Because the network's dynamics are complex and adaptive, this transformation is highly nonlinear. Finally, a simple linear model is trained to classify the original input based on these patterns.

tests, BRC achieved performance levels comparable to those of an artificial reservoir with similar dimensionality, demonstrating its potential as a viable computational paradigm.

Conclusions

This work demonstrates that living neuronal networks can serve as effective computational substrates within a reservoir computing framework. By leveraging the intrinsic dynamics of biological systems, we open a promising pathway towards neuromorphic architectures that combine energy efficiency, adaptability, and biological plausibility. Our experiments on static pattern recognition tasks confirm the feasibility of this approach and highlight its potential for future applications in AI and neuroscience. Notably, our preliminary study on this topic, presented at the ICCV 2025 Workshop "2nd Workshop on Human-inspired Computer Vision", received a Best Paper Award, and subsequent work was published at ICONIP 2025, underscoring the originality and scientific relevance of this research. Moving forward, we aim to extend BRC to more complex tasks and explore learning mechanisms within the biological reservoir, paving the way for adaptive bio-hybrid system.

References:

[1] L. Iannello et al., "From Neural Activity to Computation: Biological Reservoirs for Pattern Recognition in Digit Classification," Int. Conf. on Computer Vision (ICCV) 2025 - Workshops, pp. 4771–4780.

[2] L. Iannello et al., "From Neurons to Computation: Biological Reservoir Computing for Pattern Recognition," Neural Information Processing, pp. 114–131, Nov. 2025, doi: 10.1007/978-981-95-4100-3_9.

Please contact:

Luca Ciampi
CNR-ISTI, Italy
luca.ciampi@isti.cnr.it

Giuseppe Amato
CNR-ISTI, Italy
giuseppe.amato@isti.cnr.it

Blockchain Energy Costs for AI-Driven Scientific Infrastructure

by Enrico Barbierato and Matteo Montrucchio (Catholic University of the Sacred Heart)

Blockchains are often proposed as trustworthy backbones for AI-driven science, yet their energy costs remain poorly understood. As research infrastructures scale, these costs become a constraint rather than a footnote. This article asks what blockchain energy consumption really implies for sustainable scientific computing.

This work builds on research conducted within the project "Linea di intervento D3.2 – Scientific and Ethical Impacts of Artificial Intelligence–Based Applications", an interdisciplinary initiative of the Catholic University of the Sacred Heart in Brescia, Italy. Launched in 2019 and carried out over a 24-month period, the project examined the societal impact of Information Technology, with particular focus on finance, healthcare, and industry. Special attention was devoted to artificial intelligence techniques—especially deep learning—for large-scale data analysis, alongside issues of interpretability and ethical responsibility. Within this framework, AI systems were evaluated not only for predictive accuracy but also for broader ethical implications, anticipating the need to assess computational practices beyond performance alone.

Against this background, blockchain energy consumption emerges as a relevant case study for AI-driven scientific infrastructure. In AI-enabled research workflows, blockchains are increasingly proposed as supporting infrastructures for data provenance, auditability, reproducibility, and the secure sharing of scientific artefacts, making their computational and energy costs a non-negligible constraint. Introduced in 2008 with Bitcoin, blockchain technology operates as a distributed ledger in which transactions are grouped into cryptographically linked blocks and replicated across networks of autonomous nodes. Consensus mechanisms such as Proof of Work (PoW) require participants to perform computationally intensive tasks to validate new blocks, replacing centralised authorities with algorithmic consensus and distributing control across the network [1]. Figure 1 schematically illustrates the lifecycle of a transaction from block formation to network dissemination, consensus evaluation, and final immutability.

The decentralised nature of permissionless blockchains has made traditional financial institutions cautious, which depend on centralised governance, auditable controls, and predictable performance guarantees that conflict with the openness, volatility, and jurisdictional ambiguity of public networks. Consequently, despite growing interest in distributed ledger technologies, institutional adoption of public blockchains remains limited.

At the same time, blockchain has evolved from a niche cryptographic experiment into an infrastructure supporting cryptocurrency markets, decentralised finance, supply-chain traceability, and digital identity systems. This expansion, driven by

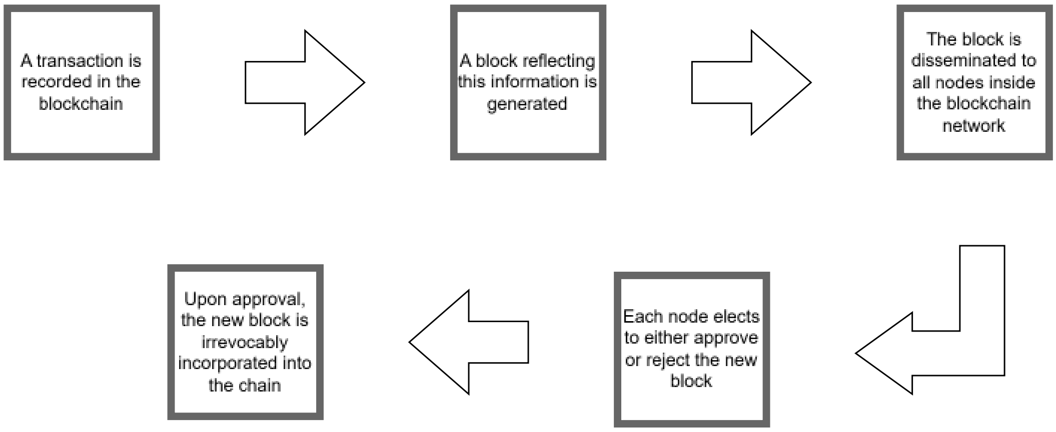


Figure 1: Flow chart of the introduction of a new block into the blockchain.

investment cycles and programmable smart-contract platforms, has also exposed structural limitations, particularly in PoW networks, where constraints in throughput, latency, and resource consumption raise concerns about long-term scalability without protocol-level changes.

A central concern is the environmental impact of PoW validation, which demands substantial computational power and, consequently, large amounts of electricity and ancillary resources such as water for cooling. Major PoW networks, most notably Bitcoin, are estimated to consume electricity on the order of tens to hundreds of terawatt-hours annually—comparable to the electricity demand of medium-sized countries. However, this cost is often mischaracterised by transaction-based metrics, as overall energy use is driven primarily by mining competition and difficulty adjustments rather than by transaction volume.

This tension between decentralization, security, and environmental impact has highlighted efficiency-oriented approaches commonly referred to as Green AI, which evaluate progress not only through predictive performance but also through computational cost, energy consumption, and resource use. By contrast, Red AI prioritises accuracy through scaling data, models, and experimentation, often with limited transparency regarding compute requirements. This distinction is particularly relevant for blockchain systems, where AI components themselves do not drive energy consumption but follow a comparable scaling logic: in PoW networks, security is effectively purchased through computation, incentivizing open-ended competition in processing power.

From a computational standpoint, transaction validation is relatively inexpensive: verifying a block containing BBB transactions requires O(B)O(B)O(B) operations, while block propagation incurs overhead that grows with network size, approximately O(N)O(N)O(N). PoW mining, by contrast, is fundamentally different, as the expected effort required to mine a block scales with the difficulty parameter DDD, yielding a cost of O(D)O(D)O(D) per block. Since DDD is dynamically adjusted to maintain a constant block interval as aggregate hashpower increases, energy consumption follows economic incentives rather than transaction volume [2].

Measuring blockchain energy use, therefore, requires a system-level perspective. For PoW networks, the most informative indicator is total electricity demand over time, typically expressed in terawatt-hours per year, as consumption is driven by aggregate hashing activity and hardware efficiency. Per-transaction or per-user metrics are often misleading, since they distribute an incentive-driven system cost over unstable or ill-defined quantities; network-level reporting, complemented by measures such as energy per block, provides a more robust basis for scientific and economic comparison.

According to the Cambridge Bitcoin Electricity Consumption Index (CBECI), maintained by the Cambridge Centre for Alternative Finance, recent annualised estimates place Bitcoin's electricity consumption in the range of approximately 90–120 TWh, with variation reflecting different modelling assumptions regarding mining hardware efficiency and profitability. Although PoW protocols impose no explicit energy cap, practical limits emerge from economic viability, hardware availability, grid capacity, regulatory constraints, and social acceptance.

Within this context, efficiency-oriented AI approaches do not seek to eliminate the intrinsic computational cost of PoW, but rather to improve measurement accuracy, transparency, and optimization under resource constraints. Data-driven models can reduce uncertainty in energy estimates, support more efficient resource management, and inform protocol-level decisions, where architectural choices often outweigh marginal efficiency gains.

Link:
https://www.nytimes.com/2018/09/24/business/walmart-blockchain-lettuce.html

References:
[1] C. Berg, Sinclair Davidson, Jason Potts – Understanding the Blockchain Economy, (Edward Elgar Publishing, 2019).
[2] Narayanan, A., Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction. Princeton University Press, 2016.

Please contact:
Enrico Barbierato
Catholic University of the Sacred Heart, Italy
enrico.barbierato@unicatt.it

Accuracy Is Not Enough: Computational Efficiency and Scientific Knowledge in AI

by Enrico Barbierato and Alice Gatti (Catholic University of the Sacred Heart)

AI models are becoming ever larger and more energy-intensive, raising questions about how scientific knowledge is produced. This article argues that computational efficiency is essential for reproducible, transparent and sustainable AI-driven science.

This contribution builds on research activities conducted within the project "Linea di intervento D.1 – A Survey about Green AI" [1], an internal research initiative of the Catholic University of the Sacred Heart in Brescia, Italy. Launched in 2024 and currently ongoing, the project examines the environmental impact of contemporary artificial intelligence algorithms and computational techniques, with a particular focus on the conditions under which AI development and deployment can be considered sustainable.

Scientific knowledge has long advanced through a structured methodological cycle combining theory formulation, experimental validation, and reproducibility. Independent verification of results is not a secondary requirement but a foundational condition that allows scientific claims to be contested, refined, and trusted. While contemporary science increasingly relies on computational modelling, data-intensive experimentation, and AI-assisted discovery, these tools must still operate within the constraints imposed by reproducibility and methodological transparency.

Since the mid-twentieth century, Artificial Intelligence (AI) has progressively reshaped scientific research by enabling large-scale data analysis, complex algorithmic modelling, and automated experimentation. More recently, Machine Learning (ML) and Large Language Models (LLMs) have accelerated discovery across domains such as climate science, biology, and physics. At the same time, they have intensified long-standing challenges related to interpretability, verification, and reproducibility [2].

Modern AI models applied to scientific problems increasingly depend on massive computational resources. Given the intrinsic complexity of many scientific domains, relatively small gains in predictive accuracy often require extensive simulations, continuous model refinement, specialised hardware, and prolonged training times. As a result, research becomes more difficult to conduct, more expensive to maintain, and harder to reproduce in practice.

Many contemporary models require dedicated accelerators, extremely long training cycles, or access to multi-million-euro computing infrastructures. Under these conditions, reproducibility is no longer guaranteed by methodological design alone, but becomes contingent on access to resources. Although scientific research is conducted primarily by academic institutions and industrial laboratories, the computational means required to replicate large-scale AI experiments are unevenly distributed. Findings and models that entail high costs, logistical complexity, or intensive energy use therefore risk being irreproducible in practice, undermining scientific rigour.

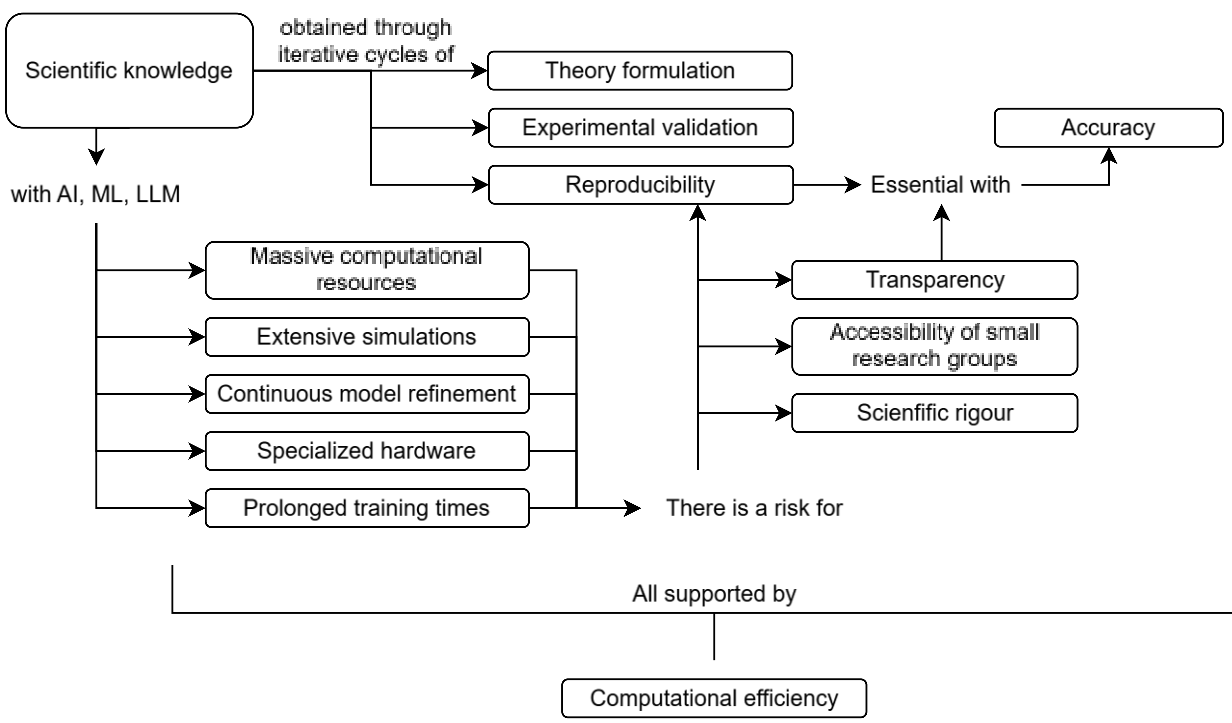


Figure 1: Scientific knowledge, AI and computational efficiency.

This issue is particularly acute in fields where computational demands scale rapidly with model complexity. In genomics and bioinformatics, high-performance computing is routinely used to analyse entire genomes and identify genetic patterns relevant to medical research. Similarly, drug discovery increasingly relies on molecular simulations and ML models that may require thousands of GPU-hours. In such contexts, reproducibility is essential, as scientific conclusions directly affect human health and safety.

Accuracy alone, however, is not a sufficient scientific metric. In many tasks, marginal improvements in predictive performance yield diminishing epistemic returns: for example, increasing accuracy from 91% to 93% may require orders of magnitude more parameters while offering little additional scientific insight. Such gains often improve numerical fit without clarifying underlying mechanisms, causal relationships, or theoretical structure. Highly optimised models thus risk becoming opaque correlational devices rather than explanatory tools.

Scientific inquiry is inherently iterative, relying on rapid cycles of hypothesis formulation, experimentation, and revision. When each iteration requires extreme computational effort, experimentation slows, alternative hypotheses are less explored, and the scientific method itself becomes constrained by a computational bottleneck. Moreover, very large models are harder to analyse and validate, more prone to hidden instabilities, and less transparent in their uncertainty, increasing the risk of undetected errors.

For AI to function as a genuinely scientific instrument, evaluation criteria must extend beyond predictive accuracy. Scientific AI systems must be stable, verifiable, analyzable, and transparent, enabling independent replication and critical scrutiny. As shown in Figure 1, computational efficiency plays a central role in supporting these classical scientific virtues: it enables faster experimentation cycles, broader accessibility for smaller laboratories, and greater robustness through simpler, more tractable models. When efficiency is treated as a first-class scientific requirement, reduced energy consumption and environmental impact follow naturally. In this sense, efficiency is an epistemic virtue of AI for science, while sustainability is its systemic by-product.

Links:
https://rescience.github.io/
https://pmc.ncbi.nlm.nih.gov/articles/PMC5016202/

References:
[1] E. Barbierato and A. Gatti, "Toward Green AI: A Methodological Survey of the Scientific Literature," in IEEE Access, vol. 12, pp. 23989-24013, 2024, doi: 10.1109/ACCESS.2024.3360705.
[2] E. Barbierato and A. Gatti, The Challenges of Machine Learning: A Critical Review. Electronics 2024, 13, 416. https://doi.org/10.3390/electronics13020416.

Please contact:
Enrico Barbierato
Catholic University of the Sacred Heart, Italy
enrico.barbierato@unicatt.it

Lightweight On-Device AI as Scientific Infrastructure for Universally Designed Augmented Reality Research

by Attila Bekkvik Szentirmai (University of South-Eastern Norway)

A browser-based augmented reality research platform demonstrates how lightweight, on-device AI can function as scientific infrastructure. By lowering technical and ethical barriers, it enables fast, privacy-preserving experimentation with computational sensing in real-world settings and across diverse user groups.

Limitations of AR Tools for Universal Design Studies

Augmented reality (AR) offers new opportunities for studying how people perceive, comprehend, navigate, and interact in space [1][2]. For accessibility and universal design (UD) research in particular, the combination of computer vision, contextual awareness, and multimodal information delivery, including audio, visual, and tactile feedback, makes AR a promising platform for investigating how digital systems can adapt to users with diverse abilities, including blind and low-vision users [3].

Despite this potential, many existing AR tools are poorly suited for use as scientific instruments in accessibility or UD studies. Research-oriented AR development often depends on proprietary hardware, AR-specific development kits, and cloud-based services. These requirements increase costs and development time, limit reproducibility, and make it difficult to involve participants early. They also raise ethical concerns when video from personal environments must be transmitted to or stored by third parties. Together, these constraints slow scientific iteration and restrict studies in everyday settings.

SensAI as Scientific Infrastructure

The SensAI project [L1], developed at Universal Design Studio USN–Bø in 2025, addresses these challenges by demonstrating how lightweight, off-the-shelf AI models can serve as dependable components in scientific research workflows. Rather than introducing new AI methods, SensAI focuses on using existing models as infrastructure that supports rapid experimentation, field-based studies, and participatory research.

The prototype is designed for researchers who need fast and reliable tools for longitudinal studies, co-design sessions, and evaluations outside laboratory environments. SensAI enables computational sensing and multimodal information delivery to be tested directly in situ, without requiring software installation or transmission of camera data to external servers. This reduces overhead while also simplifying ethical approval processes.

What the System Does

The SensAI prototype runs entirely in modern web browsers using TensorFlow.js [L2] and the COCO-SSD [L3] object detection model. As the device camera captures the user's envi-

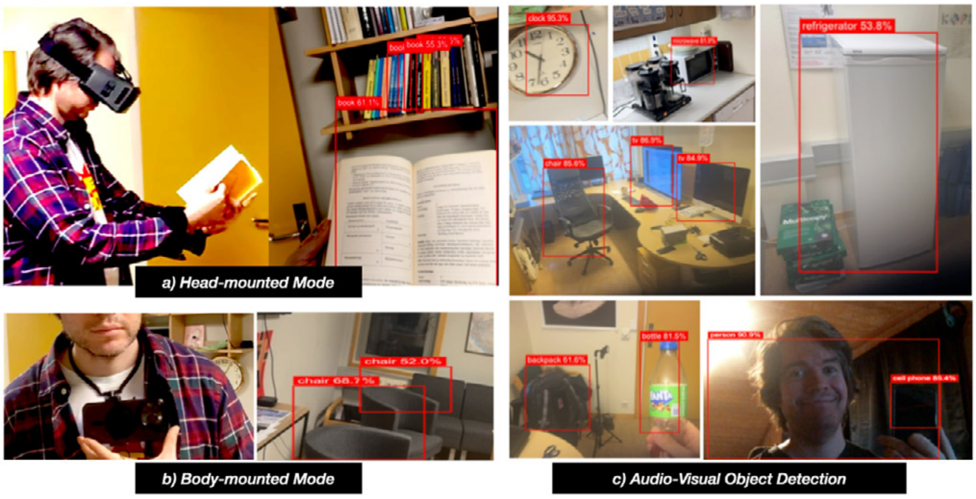


Figure 1: AR and body-mounted user interaction modalities: (A) head-mounted mode, (B) body-mounted mode, and (C) multimodal audio–visual object detection.

ronment, the system identifies and visualises objects while providing short spoken cues, with an optional visual overlay (see Fig. 1c).

The feedback style follows the synthetic speech commonly used by screen readers. This supports accessibility while allowing the same system to be used by participants with different sensory abilities. Rather than generating long descriptions, SensAI delivers brief, audio-centred object cues. This reframes AR information delivery beyond visual-first paradigms and supports research into alternative perceptual strategies.

Interaction Modalities

The prototype supports several interaction modalities:

- Handheld smartphone mode
- Body-mounted, hands-free mode with low physical effort (see Fig. 1b)
- Immersive head-mounted AR passthrough mode using cardboard accessories (see Fig. 1a)
- Stationary mode using a desktop or laptop webcam, or a tablet.

This flexibility reflects UD principles by accommodating different comfort levels and preferences for movement and device use. For researchers, it enables direct comparison of interaction styles across contexts using the same system, without building multiple prototypes.

Why On-Device AI Matters for Scientific Research

One of SensAI's key contributions is showing how small, on-device AI models can support scientific AR studies. Many AR research platforms rely on proprietary hardware and cloud-based processing. While powerful, these approaches increase development time, require specialised expertise, and complicate data management and ethical review.

By running COCO-SSD directly in the browser, the prototype offers properties that are directly relevant to scientific practice:

- Real-time audiovisual cues that support studies of perception and interaction
- Offline use once the model is loaded, enabling fieldwork in varied environments
- Strong privacy protection, since no images or personal information leave the device
- Instant deployment and replication through a simple URL share.

Together, these characteristics make the system suitable for exploratory studies, comparative experiments, and reproducible research beyond controlled laboratory settings.

Implications for AI for Science

SensAI illustrates how AI can function as scientific infrastructure rather than as a standalone application. Lightweight, browser-native AI components allow researchers to focus on study design and scientific questions instead of system integration and deployment logistics. This approach is particularly relevant for accessibility and UD research, where early and repeated participant involvement is essential. More broadly, this work shows that progress in AI for science does not always depend on large-scale or cloud-based systems. Carefully chosen, openly available AI components can support trustworthy and reproducible research across disciplines.

Links:

[L1] SensAI: https://doi.org/10.5281/zenodo.17931940

[L2] TensorFlow.js: https://www.tensorflow.org/js

[L3] https://kwz.me/hGJ

References:

[1] M. Billinghurst, A. Clark, and G. Lee, "A survey of augmented reality," Foundations and Trends® in Human–Computer Interaction, vol. 8, no. 2–3, pp. 73–272, 2015.

[2] A. B. Szentirmai, "Universally designed augmented reality as interface for artificial intelligence assisted decision-making in everyday life scenarios," Studies in Health Technology and Informatics, vol. 320, pp. 469–476, 2024.

[3] A. B. Szentirmai, "Enhancing accessible reading for all with universally designed augmented reality: AReader," in Proc. Int. Conf. on Human-Computer Interaction, pp. 282–300, Springer Nature, 2024.

Please contact:

Attila Bekkvik Szentirmai, University of South-Eastern Norway, Department of Business and IT, Norway
attila.b.szentirmai@usn.no

What Promises do AI Hold for Computational Imaging?

by Tristan van Leeuwen, Felix Lucka and Ezgi Demircan-Tureyen (CWI)

An image says more than a 1000 thousand words, it is said. This also holds true in many scientific applications, where 2D, 3D, or even 4D images are analysed. But how do we compute images from raw measurements, and can AI help us improve? At CWI's Computational Imaging group in Amsterdam, mathematicians and computer scientists are trying to answer these questions.

Imaging is a key tool in many sciences. Physicists and material scientists use electron microscopes to study nanoparticles, biologists study organoids under a microscope. At much larger spatial scales, earth scientists use earthquake data to image the deep earth and astrophysicists peer into deep space using telescopes. What unites these fields is that they aim to image things that are not directly observable with the naked eye, or even inaccessible to humans. Advances in instrumentation have revealed things at increasingly smaller scales and at larger distances. Still, direct observation has its limitations – however powerful the instrument – as observations are not always directly related to the features of interest and may contain noise or other imperfections.

To get an interpretable image from the measured data, we need to solve an inverse problem. There, we fit a mathematical model of the image formation process to the measured data. A well-known example is image deblurring, where the imprint of the optical system on the image needs to be removed. A medical CT scan is another well-known example. Here X-ray radiographs are taken from all around the patient and are combined into a 3D reconstruction of the patient's inner anatomy. In radioastronomy, a similar situation arises and measurements can be processed into a 2D image of part of the sky at a certain distance.

In some cases, when enough measurements of high quality are available, the inverse problem can be solved explicitly and the resulting formulas implemented in software to yield an efficient algorithm that processes the measured data into an image. More advanced mathematical image reconstruction techniques can work with less data and handle more noise, but they require something in return [1]. The missing information needs to be added implicitly or explicitly by making prior assumptions about the object we aim to image. These methods come with two major challenges, however: i) the computational cost of advanced image reconstruction methods is too high for many practical applications; and ii) it is difficult to capture suitable prior information in hand-crafted mathematical models.

So how can AI help us address these challenges? For one, AI can reduce the computational cost of advanced image reconstruction algorithms by replacing an expensive iterative process by a single pass through a deep neural network. The computational cost is then shifted from solving each problem individually to a one-time training phase. This comes in different flavours [2]. Learned inverses treat image reconstruction as pattern recognition, learning to map measurements to images in a single pass. Unrolling integrates physics into the neural network, compressing many iterations into fewer physics-aware steps that reach good solutions faster. Recently, diffusion models have pushed the idea further and allow one to alternate between generating plausible images and fitting the measured data. In our group we are currently working on ways to incorporate physics and measured data into such generative models. Just as you might prompt a text-to-image model, we could then prompt it with "Show me the 3D volume of the chest that could have produced these X-ray radiographs."

To effectively deploy AI for scientific imaging, though, representative training data are needed. The most impressive AI image generators were trained on billions of images collected from the internet, that are unfortunately not representative for

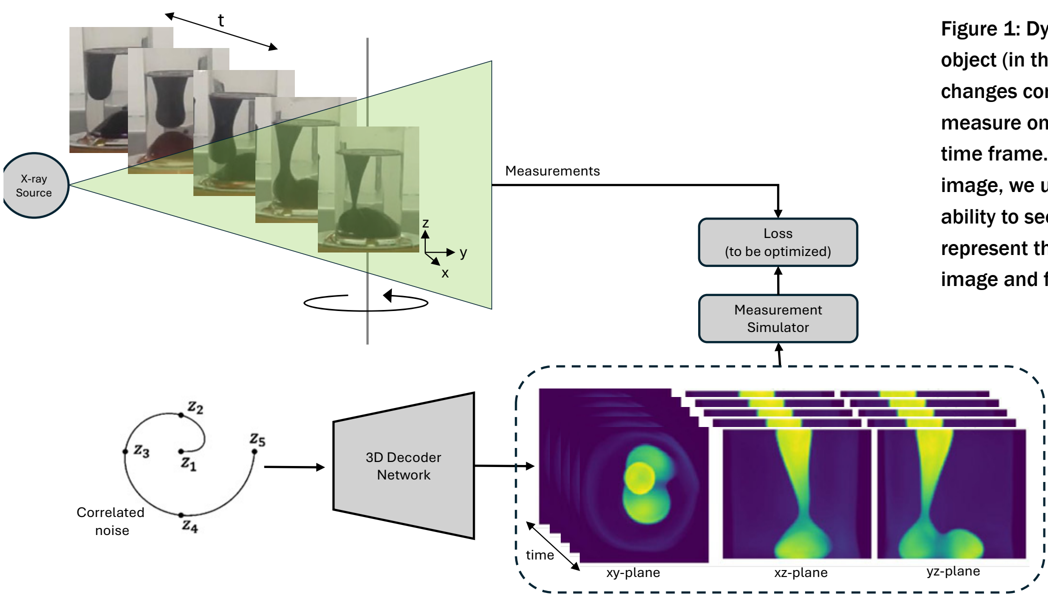


Figure 1: Dynamic CT imaging: the object (in this case, a lava lamp) changes continuously, but we can measure only one X-ray radiograph per time frame. To reconstruct the 4D image, we use a neural network's ability to see structure in noise to represent the dynamically evolving image and fit it to the measured data.

most scientific imaging applications. CWI's Flex-ray laboratory [L1] is an experimental imaging lab which was founded for the purpose of developing and validating computational imaging algorithms. We use our in-house CT scanner, amongst other things, to collect large datasets specifically for the purpose of testing and validating AI for scientific imaging [3].

Collecting large data sets comprising measurement data and matching images is difficult and costly, and in medical imaging, further complicated by privacy concerns. In some applications, therefore, we do not have pairs of measurement data and matching ground-truth images. For these cases, we are also actively developing un- or self-supervised training paradigms. In some cases, training data may not be available at all. This happens for example in 4D CT where the object dynamically evolves over time while it is being scanned. A static 3D snapshot of the ground truth needed for supervised training can therefore never be obtained. Surprisingly, neural networks can still help [3]. An example of our recent work is shown in Figure 1, where we use a neural network's ability to see structure in noise to represent the dynamically evolving image and fit it to the measured data.

In summary, AI can strengthen computational imaging in several ways. It can be used to speed up the image reconstruction process, promising near real-time performance for specific tasks. Moreover, generative models can capture structure in images in an unprecedented way. It is important, however, that the use of AI in science stays grounded in mathematics and is aware of the underlying physics. The motto should be: model what we can and learn what we must.

Link:
[L1] https://www.cwi.nl/en/collaboration/labs/flex-ray-lab/

References:

[1] S. Ravishankar, J. C. Ye, and J. A. Fessler, "Image reconstruction: From sparsity to data-adaptive methods and machine learning," Proc. IEEE, vol. 108, no. 1, pp. 86–109, Jan. 2020.
[2] S. Arridge, P. Maass, O. Öktem, and C. B. Schönlieb, "Solving inverse problems using data-driven models," Acta Numerica, vol. 28, pp. 1–174, 2019.
[3] M. B. Kiss, et al., "2DeteCT—A large 2D expandable, trainable, experimental computed tomography dataset for machine learning," Sci. Data, vol. 10, no. 1, Art. no. 576, 2023.

Please contact:

Tristan van Leeuwen, CWI, The Netherlands
t.van.leeuwen@cwi.nl

Detecting Small Changes in 3D Aerial Scans: A Hybrid Approach for Real Urban Environments

by Tatjana Ćeranić, Stephan Schraml and Philip Taupe (AIT Austrian Institute of Technology GmbH)

Reliable detection of small, local changes in real airborne laser scanning data remains difficult with current 3D change-detection techniques. Off-the-shelf methods often overlook subtle modifications or flag too many false positives. By combining semantics with geometry-based and deep-learning methods, we aim to improve robustness in noisy, cluttered settings.

Airborne laser scanning (ALS) is increasingly used to capture detailed 3D representations of complex outdoor environments. Beyond large-scale urban mapping, ALS enables the monitoring of subtle, localised physical changes, e.g., the sudden appearance or removal of smaller structures or objects – changes that can carry operational or safety implications, particularly in sensitive areas. Detecting these kinds of changes is far more challenging than identifying major modifications such as new buildings or terrain reshaping, yet it is precisely this fine-grained analysis that many real-world applications require.

However, reliably identifying such objects in the highly variable conditions of real ALS acquisitions remains difficult. The changes of interest are typically less than two metres in size, sometimes partially occluded, and recorded under different flight paths. At the same time, the surrounding geometry, such as vegetation, clutter, or naturally irregular surfaces, introduces significant noise and variability. For applications where even a seemingly minor modification can matter, existing 3D change-detection methods do not yet offer the robustness needed in practice.

Why existing methods struggle with real ALS data

Deep-learning methods for 3D change detection are typically trained using publicly available datasets, most of which are synthetic and therefore only approximate real airborne laser-scanning (ALS) conditions. They represent idealised versions of the data: noise-free, uniformly sampled, free of occlusion artefacts, and lacking the complex density fluctuations and natural clutter present in real operational settings. In principle, a dedicated real-world training set could overcome these differences, but assembling one that reflects a specific sensor setup, acquisition procedure, and the range of environmental variability (e.g. seasonal changes) requires substantial effort in data collection and manual annotation. For many applications, creating such a dataset is simply not feasible. Consequently, a domain gap emerges between the synthetic training data and real ALS acquisitions, and models trained on synthetic datasets may generalise less well under real operational conditions, for example by detecting too many false positives.

Classical geometric change-detection techniques (e.g. difference of DEMs, cloud-to-cloud comparison, M3C2) may seem

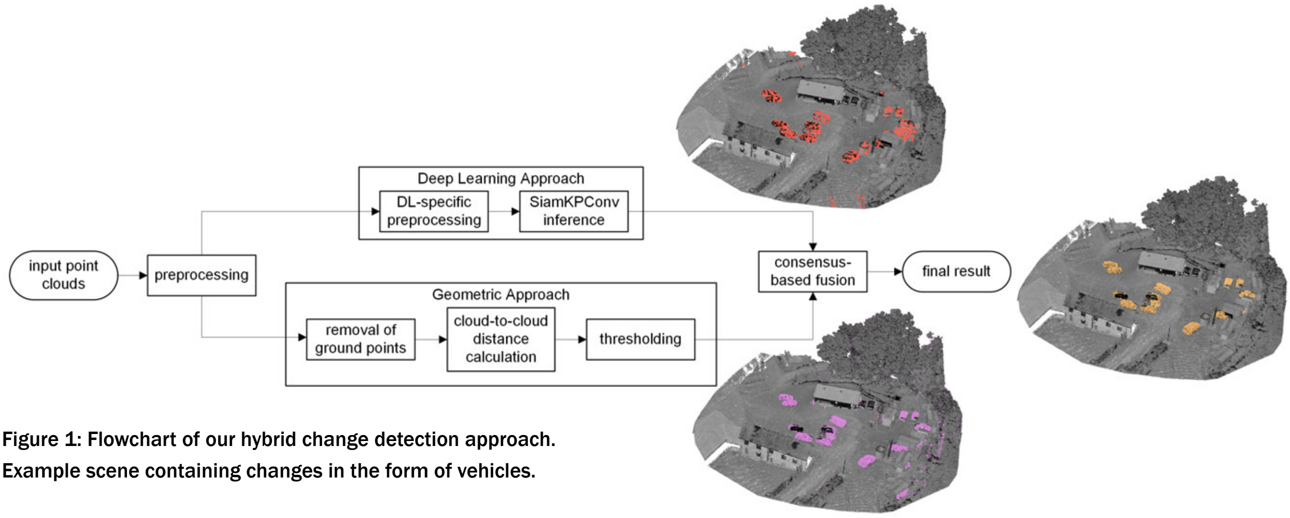


Figure 1: Flowchart of our hybrid change detection approach. Example scene containing changes in the form of vehicles.

like a robust alternative that does not rely on large amounts of training data, but they also face challenges. In environments with thin or delicate structures (e.g. power lines, flags), dense vegetation, or irregular surfaces, these methods can produce a high number of false detections because they interpret natural variability as structural change.

A hybrid strategy tailored to real conditions

To address these limitations, we developed a hybrid change-detection workflow which combines the strengths of geometry-based analysis with the flexibility of deep learning. Crucially, it functions entirely with a deep-learning model trained on openly available data and requires no additional training data acquisition.

The workflow begins with preprocessing tailored to real ALS conditions. Both point clouds are pre-processed to ensure comparable characteristics (density, denoising, etc.) before the workflow branches into two complementary detection methods:

1. A geometric, learning-free change detector: The point clouds are separated into ground and non-ground points using a cloth simulation filter [1]. Only the non-ground points are analysed further, as the focus is on detecting structural or object-level changes rather than terrain fluctuations. Using nearest-neighbour distances between the two surveys, the system flags points that have shifted significantly compared to the reference point cloud. This method excels at picking up clear geometric differences.
2. A deep-learning-based detector using a Siamese KPConv network (SiamKPConv) [2,3]:

We use a variant of the SiamKPConv originally developed for urban change detection, trained on the corresponding synthetic dataset. Although the network was trained on a data distribution different from the real-world ALS data, its learned structural features still provide valuable cues for identifying regions likely to represent real physical changes. The SiamKPConv model also outputs semantic labels for detected changes. Although not always fully accurate for ALS data, these labels provide helpful hints about whether a change resembles vegetation, built structures, or mobile objects, thus supporting post-processing filtering and quicker interpretation of unexpected modifications.

Independently, each branch produces its own set of changed point candidates. To reduce false positives, these results are merged using consensus-based fusion: a point flagged by the geometric method must lie close to a corresponding detection from the deep-learning method. This dual confirmation significantly improves reliability in noisy or cluttered scenes, as it ensures that only those changes supported by both geometric evidence and learned features are kept, while retaining the higher resolution provided by the geometric method (see Figure 1; note how the fusion step suppresses vegetation-related noise).

What the system can detect

Our current focus is on urban objects such as street furniture, temporary barriers and general clutter. Larger changes, such as new buildings or significant adjustments to existing infrastructure, are also captured reliably. Although detecting vegetation changes is not the system's primary focus, they are detected to some extent through the deep-learning branch. The semantic labels produced by the SiamKPConv network can further support analysis by suggesting which type of object might have changed.

Outlook

Possible future developments include integrating object-level grouping (e.g. connected components, instance segmentation) to produce map-ready outputs, introducing more advanced semantic descriptors to target specific types of changes. Furthermore, exploring semi-supervised or unsupervised strategies to gradually adapt the deep-learning model to real ALS characteristics could help to bridge the domain gap.

References:

[1] W. Zhang, et al., "An easy-to-use airborne LiDAR data filtering method based on cloth simulation," Remote Sens., vol. 8, no. 6, Art. no. 501, 2016.
[2] I. de Gélis, S. Lefèvre, and T. Corpetti, "Siamese KPConv: 3D multiple change detection from raw point clouds using deep learning," ISPRS J. Photogramm. Remote Sens., vol. 197, pp. 274–291, 2023.
[3] I. de Gélis, T. Corpetti, and S. Lefèvre, "Change detection needs change information: Improving deep 3-D point cloud change detection," IEEE Trans. Geosci. Remote Sens., vol. 62, pp. 1–10, 2024.

Please contact:

Philip Taupe
AIT Austrian Institute of Technology GmbH, Austria
philip.taupe@ait.ac.at

Fully Automated Detection of Harmful Cyanobacteria Blooms in Lakes Using Photo Traps and Machine Learning

by Jean-Baptiste Burnet and Olivier Parisot (LIST)

Cyanobacteria blooms pose growing risks to drinking water supplies and recreational waters, a challenge intensified by climate change and inadequately captured by current regulatory monitoring strategies. Our study demonstrates how low-cost, ground-based RGB cameras combined with machine learning enable near real-time detection of blooms. By deploying automated photo traps and YOLO-based detection models at a major freshwater reservoir in Luxembourg, we open new pathways for early warning systems and improved understanding of harmful cyanobacteria bloom dynamics.

Cyanobacteria (also called blue-green algae) blooms in freshwater supplies worldwide threaten the health of bathers, domestic animals and livestock due to the toxins released by these microorganisms. Fuelled by climate change, this phenomenon further undermines the environmental health of aquatic ecosystems. The highly fluctuating dynamics of harmful cyanobacteria blooms (CyanoHABs) require representative monitoring frameworks that are hardly achieved using regulatory monitoring strategies currently in place. New tools are therefore needed to enhance our understanding of bloom dynamics and provide early warning systems to proactively react to such events, both for safe drinking water supply and for beach management [1]. Ground-based remote sensing is an interesting alternative to satellite imagery, which is limited by cloud cover and revisit times.

At LIST, we have implemented a fully automated in situ image acquisition workflow to detect blooms at near real-time (Fig. 1). The study was performed on the Upper-Sûre Lake, the main drinking water supply in Luxembourg and a major recreational area that experiences cyanobacteria blooms every year in late summer. Two sites were equipped since 2021 with commercial photo traps (Hyperfire 2, Reconyx). Pictures (in RGB) were taken hourly from June-November and sent in real-time to the researchers at LIST using the standard simple mail transfer protocol (SMTP).

For automatic bloom detection using machine learning, we adapted the classic object detection process. The first phase consisted of creating an annotated dataset containing images and the position of blooms (in the form of bounding boxes). The original images came from different stations at different times of the year and were then cut into 640 x 640 patches. Based on our expertise in the field, we then meticulously edited the bounding boxes using MakeSense software [L1]. In the end, we preselected a set of 1,061 images, which were then divided as follows: 599 images for training, 181 images for validation, and 281 images for testing. We aimed to retain enough ‘background’ images for validation and testing to reduce the occurrence of obvious false positives, particularly in cases of sunlight reflecting off water, mirror effects with trees,

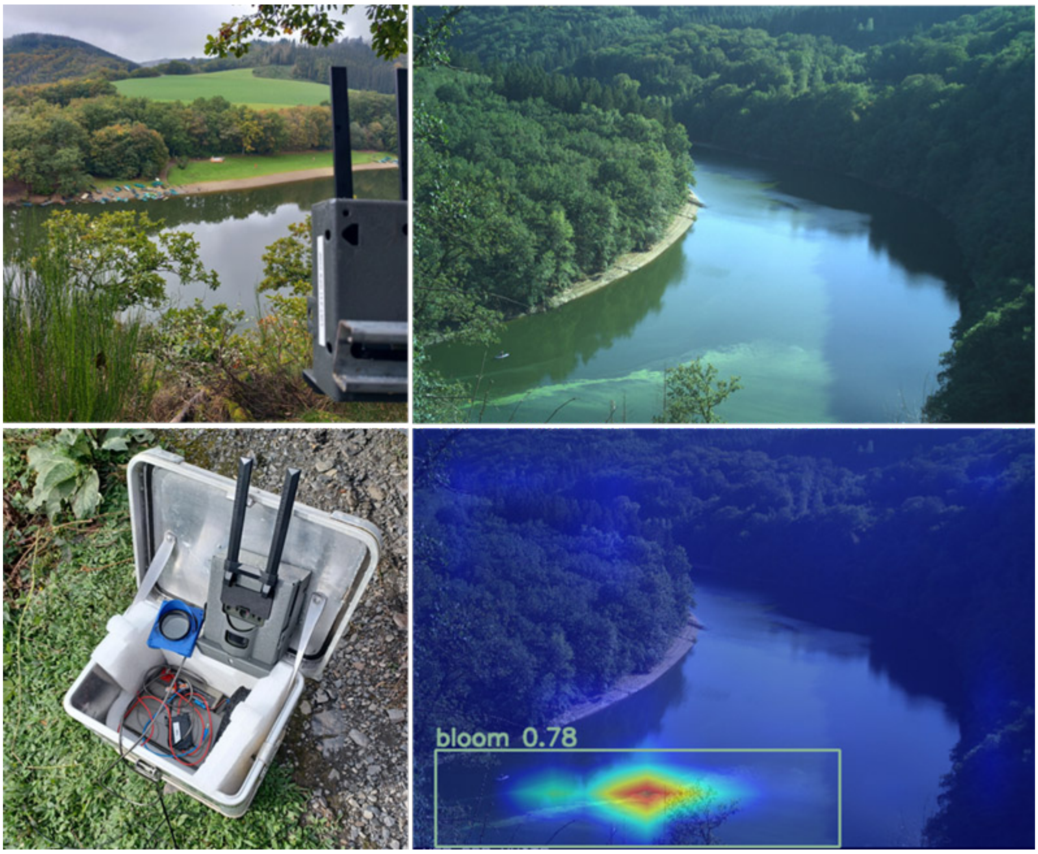


Figure 1: Harmful cyanobacteria blooms are detected in the Upper-Sûre Lake, Luxembourg, in near real-time using machine learning combined with ground-based remote sensing.
Credits: LIST.

etc. The second phase consisted of training a detection model based on this data, using the YOLO (You Only Look Once) architecture. There are different versions of YOLO, but we chose version 7 [2] because it offers a highly effective training and evaluation pipeline implementation [L1], as well as a software licence that is standard in an academic context (GPLv3). The annotated images were used to train a YOLOv7 model by applying transfer learning, i.e., using the default pre-trained YOLOv7 model. Finally, we performed multiple model trainings (500 epochs, different models' sizes, with or without training-time data augmentation). Considering the overall metrics, the best-performing model relies on a tiny architecture (7M parameters; precision = 0.531, recall = 0.352, mAP50 = 0.325). However, it is worth noting that two models based on a normal architecture (37M parameters) achieve either higher precision (0.623) or higher recall (0.461). We then used Grad-CAM [L2], an explainable AI technique, to visualise the regions of the image that most strongly influence the model's predictions, providing valuable insights into false positive and false negative cases. Other alternatives are currently being evaluated, including more recent versions of YOLO (8, 11, 12), RET-DETR (Vision Transformer-based real-time object detector), as well as older ones such as FasterRCNN (Real-Time Object Detection with Region Proposal Networks).

In 2024 and 2025, additional photo traps have been equipped with transmission systems and thousands of pictures have been taken during both seasons and analysed manually by LIST researchers. They will now serve to further train the selected YOLO models and improve their performance. Overall, these valuable datasets enable a better understanding of the fine spatial and temporal dynamics of cyanobacteria blooms and will constitute crucial inputs for the development of forecasting models.

Links:
[L1] https://www.makesense.ai/
[L2] https://github.com/jacobgil/pytorch-grad-cam

References:
[1] H. Almuhtaram, et al., "State of knowledge on early warning tools for cyanobacteria detection," Ecol. Indic., vol. 133, Art. no. 108442, 2021.
[2] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," arXiv preprint, arXiv:2207.02696, 2022.

Please contact:
Jean-Baptiste Burnet, Luxembourg Institute of Science and Technology (LIST), Luxembourg
jeanbaptiste.burnet@list.lu

AI-enhanced Operational Picture for Public Safety Operations from Heterogeneous Information Sources

by Julia Pöschl, Philip Taupe, Jakob Hurst (AIT Austrian Institute of Technology GmbH)

Staying informed is crucial for decision-makers, particularly in time-critical domains such as disaster response and public safety. The Enhanced Language Interpreter (ELI) presented here supports decision-makers in making use of information that is available but cumbersome to process. It analyses heterogeneous inputs using large language models and knowledge graphs and converts them into a concise representation tailored to the target domain, thereby enhancing the operational picture of the situation.

Today's age of information and communication technology offers the possibility to receive information updates on almost everything that is going on anywhere around the world in near real time. This abundance of information can be leveraged to enhance the operational picture of e.g., emergency forces in disaster response or other civil protection scenarios. The palette of information sources ranges from social media, news outlets, and (space-based) earth observation services, to in-field reports of front-line responders. The abundance of information brings a major challenge: extracting domain-relevant material, filtering out noise, and delivering an accurate operational picture. This challenge is further complicated by the inhomogeneity of data sources with respect to the used domain-specific vocabulary and the way information is presented, which makes it challenging to establish one coherent operational picture to support decision-making.

To alleviate these challenges, we develop ELI (Enhanced Language Interpreter), a powerful AI-based tool based on principles of semantic cross-domain interoperability to extract, categorize, interlink, and present information to decision makers. The research incentive behind ELI is to investigate AI-driven systems to improve situational awareness in complex emergency response settings, safety-related operations, and similar contexts. The tool evolved from a series of national and international research and development projects starting from 2020 and is being developed by the Austrian Institute of Technology GmbH.

ELI focuses on textual input and operates on the assumption that any information input can be translated into a written form and hence ingests unstructured or semi-structured incident reports with varying degrees of detail, or any other textual data gathered from heterogeneous sources. ELI offers a concise, structured representation of knowledge in the form of relevant entities and their interrelationships, as well as the possibility to trace these back to their individual occurrences in the source documents. One of the main strengths of our tool is its capability to align its output with domain-specific vocabulary and

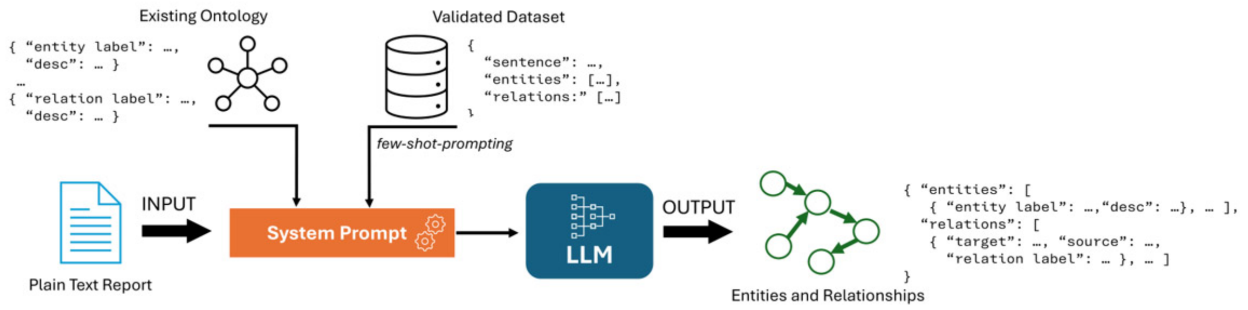


Figure 1: ELI's information extraction and categorization process utilizing large language models.

ontologies, tailored to the specific needs of decision makers in various domains.

ELI leverages natural language processing and knowledge graph techniques for gaining insights on written documents. In a first step (Figure 1), the system utilizes state-of-the-art large language models (LLMs) to identify entities and their relationships in a text. This step also includes categorizing the extracted information according to a pre-supplied domain-specific target ontology. In a second processing step, links between entities and groupings thereof are identified across all documents using semantic similarity matching. Finally, the system provides a graph representation of the extracted information including references into the source documents. A graph-RAG [1] based module is currently under active development, allowing for a chatbot-like interaction between ELI and the user to access the gathered knowledge.

Multiple evaluations were performed to investigate ELI's applicability beyond specific safety or security missions. The underlying methods were tested against common benchmarks and synthetic datasets. This includes GraphRAG-Benchmark [2], a domain-specific question-answering benchmark for information retrieval and reasoning on graphs in fiction and medical domain. RE-3d [L1], a knowledge graph construction dataset in the defense domain, was used to validate the information extraction process. Finally, ELI's overall approach has been verified with stakeholders in a controlled live demonstration.

As ELI is currently in the stage of a first prototype, there are many possible future directions for research and development. Most obvious is the improvement and expansion of its methods in terms of utility, accuracy and efficiency, for example considering other entity grouping strategies or missing link prediction. Moreover, we hope to be able to validate the approach in real-world scenarios with stakeholders from different domains in the future. Widening the scope of ELI, one could also further investigate other types of input data, e.g. by adding speech transcription and image description generation, or testing the tool against them.

To summarize, most of the information that emergency responders and other decision makers work with comes in or can be represented in a written form. ELI is a tool designed to help gain situational awareness for decision support by efficiently analyzing such written information. It leverages large language models, semantic similarity search, and knowledge graphs to structure information into an operational picture. With its semantic cross-domain interoperability credo, it is especially capable to handle inputs with heterogeneous forms and domain vocabulary.

Link:
[L1] https://github.com/dstl/re3d

References:

[1] H. Haoyu et al., "RAG vs. GraphRAG: A Systematic Evaluation and Key Insights", 2025, arXiv:2502.11371.
[2] Z. Xiang et al., "When to use Graphs in RAG: A Comprehensive Analysis for Graph Retrieval-Augmented Generation", 2025, arXiv:2506.05690.

Please contact:

Jakob Hurst
AIT Austrian Institute of Technology GmbH, Austria
jakob.hurst@ait.ac.at

Context-Guided Evolutionary Algorithms for Consensus Inference of Gene Regulatory Networks

by Adrián Segura Ortiz, José García-Nieto and Ismael Navas Delgado (ITIS Software, University of Málaga)

The resulting consensus networks are more robust and interpretable, enabling deeper insights into complex disease mechanisms.

Understanding how genes interact with each other is essential for explaining how cells function and why diseases develop. Advances in molecular biology now enable us to measure thousands of genes simultaneously. However, transforming these measurements into accurate maps of regulatory interactions remains a significant challenge. Different computational algorithms tend to disagree, and many overlook well-established biological principles, resulting in models that are mathematically correct but biologically implausible.

Responding to this need, researchers at the University of Málaga (Spain), within the Khaos Research Group [L1] and the ITIS Software Institute [L2], have developed a framework for context-guided consensus inference of gene regulatory networks (GRNs). Instead of relying on a single method, the project combines multiple existing techniques and aligns them with biological knowledge. Individual algorithms are treated as complementary “opinions”, and consensus networks are produced through evolutionary optimisation guided by biological realism rather than purely mathematical scores (Figure 1).

The initiative began in 2022 as part of a doctoral thesis and progressed through a sequence of increasingly biologically guided tools:
- GENECI [1], the first release, demonstrated that evolutionary search could combine predictions from multiple GRN inference techniques. Using a simple genetic algorithm, it produced more stable networks than individual methods alone, showing that consensus could be treated as an optimisation problem rather than a voting procedure.
- Memetic Inference extended this idea by incorporating a small number of known gene–gene interactions into the optimisation process. A local search component helped refine the networks towards these known links, improving accuracy without overfitting to prior knowledge.
- MO-GENECI [2] then shifted from a single optimisation goal to a multi-objective interpretation, incorporating structural and functional properties often observed in real regulatory systems, such as characteristic topologies and regulatory motifs. This allowed the algorithm to preserve agreement among methods while also favouring biologically plausible organisation.
- PBEvoGen brought domain experts into the loop. Rather than merely interpreting results, specialists could steer the optimisation toward regions of biological interest, transforming consensus inference into an interactive process guided by expert knowledge.
- Finally, BIO-INSIGHT [3] generalised this approach into a many-perspective strategy, simultaneously evaluating regulatory networks from several biological viewpoints (including expected interaction densities, dynamic behaviour, motif enrichment, and reduction of spurious links) and scaling efficiently to large datasets through distributed computation.

These tools are not isolated prototypes. They have been unified in GENECI, an open-source Python package [L3] that integrates the orchestrator of more than 25 GRN inference techniques with the full evolutionary framework. The package is freely available to the community [L4] and has already been downloaded over 16,000 times, demonstrating interest from research groups working in gene regulation, computational biology, and clinical data mining.

This progression has been driven by a practical need: in biomedical contexts, biological relevance takes precedence over numerical fit. A network may achieve excellent statistical performance yet fail to represent meaningful regulatory behaviour. For example, some algorithms prioritise noise reduction and lose regulatory motifs, while others perform well in yeast but fail in human tissue. By favouring biologically coherent interactions, the framework reduces false positives and improves interpretability. Its aim is to generate credible hypotheses that scientists and clinicians can meaningfully investigate.

The framework has been validated using both simulated benchmarks and real-world clinical data. Collaborators provided transcriptomic datasets from melanoma, fibromyalgia, and myalgic encephalomyelitis. In these studies, the algorithms identified regulatory patterns associated with inflammation and immune dysregulation. While not designed for di-

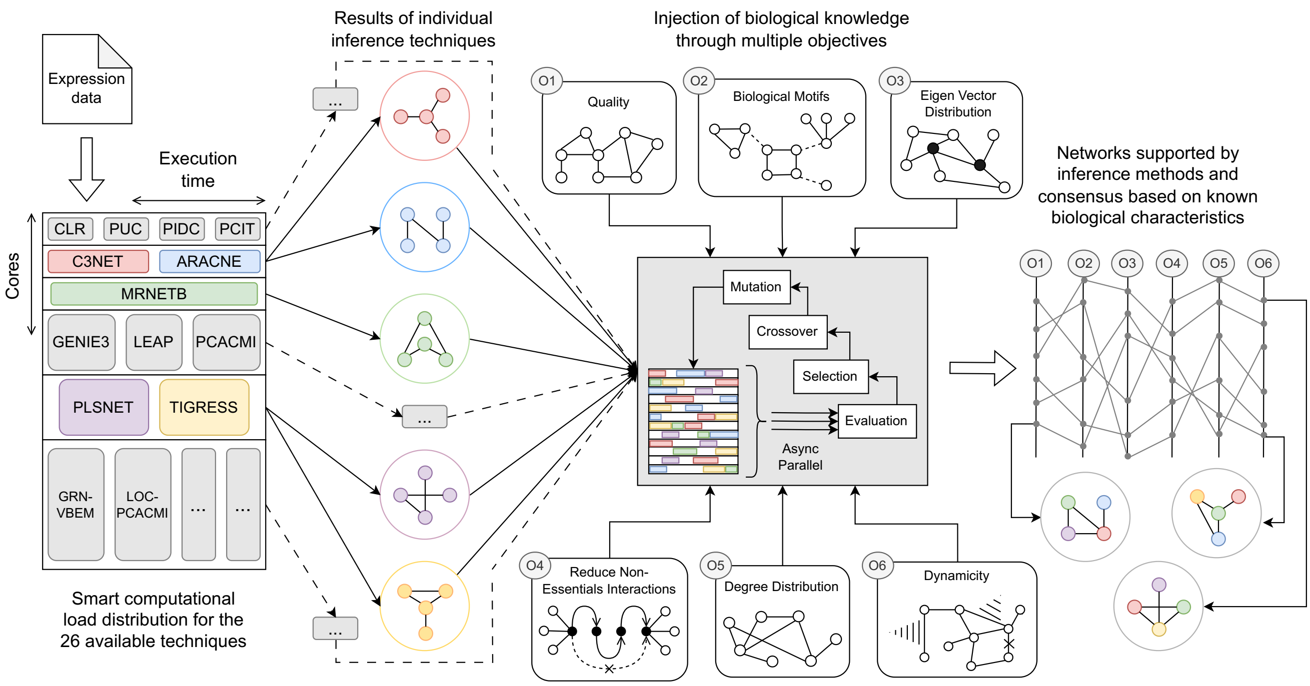


Figure 1. Overview of the context-guided consensus proposal for inferring gene regulatory networks. Multiple inference techniques are executed in parallel, each producing a different candidate network. These networks are evaluated and combined through an evolutionary process informed by biological knowledge, including regulatory patterns, interaction quality, and realistic network structures. The resulting consensus networks are both supported by the computational evidence and constrained by known biological characteristics.

agnosis, the results provide testable biological leads, supporting research into conditions where reliable biomarkers are still lacking.

Collaboration has been essential. A research stay at the University of Lille (France) contributed expertise in evolutionary computation, while clinical partners at the University of Valencia (Spain) provided access to medical data and biological interpretation. This cross-disciplinary model reinforces the project's commitment to reproducibility: all methods are openly released, documented, and designed to be extended by the scientific community.

Looking ahead, the team plans to adapt consensus inference to single-cell transcriptomics, enabling analysis of regulatory variability between individual cells. They are also developing interactive tools that allow experts to guide optimization in real-time. Together with opportunities for cooperation within the ERCIM community in regulatory modelling, explainable AI, and biomedical graph mining, these developments highlight a broader ambition: to bridge computational innovation with biological insight and deepen our understanding of the regulatory mechanisms that govern life.

Links:
[L1] https://khaos.uma.es/
[L2] https://itis.uma.es/
[L3] https://pypi.org/project/geneci/
[L4] https://github.com/AdrianSeguraOrtiz/GENECI

References:

[1] A. Segura-Ortiz, et al., "GENECI: A novel evolutionary machine learning consensus-based approach for the inference of gene regulatory networks," Comput. Biol. Med., vol. 155, Art. no. 106653, 2023.
[2] A. Segura-Ortiz, et al., "Multi-objective context-guided consensus of a massive array of techniques for the inference of gene regulatory networks," Comput. Biol. Med., vol. 179, Art. no. 108850, 2024.
[3] A. Segura-Ortiz, et al., "Multifaceted evolution focused on maximal exploitation of domain knowledge for the consensus inference of gene regulatory networks," Comput. Biol. Med., vol. 196, Art. no. 110632, 2025.

Please contact:

Adrián Segura Ortiz
ITIS Software, University of Málaga, Spain
adrianseor.99@uma.es

José Francisco Aldana Montes,
ITIS Software, University of Málaga, Spain
jfaldana@uma.es

Learning Wind-Turbine Wakes with Generative Adversarial Networks

by Jamal Toutouh (University of Málaga, Spain), Sergio Nesmachnow (Universidad de la República, Uruguay), Martín Draper (Universidad de la República, Uruguay), and Maximiliano Bove (Universidad de la República, Uruguay)

Data-driven and physics-informed generative adversarial networks provide fast surrogates for wind-turbine wakes, bridging high-fidelity simulation and wind farm design.

Wind-farm design and control require reliable predictions of how turbines interact through their wakes. Wake interactions determine power losses, mechanical loads, and fatigue in downstream turbines, strongly influencing wind-farm efficiency and lifetime . High-fidelity approaches based on Large Eddy Simulation with Actuator Line Models (LES-ALM) can resolve these flows in detail, but each simulation may require days of computation. That cost limits systematic exploration of new layouts, operating conditions, and control strategies.

Recent work in scientific machine learning has proposed physics-informed generative adversarial networks (PI-GANs) as a way to combine the strengths of first-principles models and deep generative learning [2]. PI-GANs extend the standard generative adversarial network (GAN) framework by adding loss terms that measure how well generated samples satisfy the governing physical laws. The GAN discriminator still distinguishes real and generated data, but the generator is additionally regularised to respect conservation laws or constitutive relations. In fluid mechanics, such models promise to reduce data requirements and improve extrapolation to new regimes, while retaining physical consistency.

The work presented here belongs to a broader research effort on physics-informed generative models for fluid flows in wind energy. The study is a collaboration between the Instituto de Mecánica de los Fluidos e Ingeniería Ambiental and the Instituto de Computación at Universidad de la República (Uruguay), together with the Instituto de Tecnologías e Ingeniería del Software at Universidad de Málaga (Spain) [L1]. The team has developed and analysed a purely data-driven GAN surrogate model for wind-turbine wakes, which provides a baseline against which future PI-GAN variants can be assessed [1].

The proposed surrogate model targets the mean streamwise wind speed at hub height in the wakes of a virtual wind farm with 15 turbines [1]. The dataset is generated with LES-ALM simulations of seven layouts and five inflow wind speeds. The inflow boundary condition comes from precursor simulations of an atmospheric boundary layer with realistic turbulence. For each configuration, the time-averaged velocity field on a horizontal plane at hub height is stored.

Each training sample consists of two elements: the horizontal wind profile two rotor diameters upstream of a selected turbine, and the corresponding mean downstream velocity field

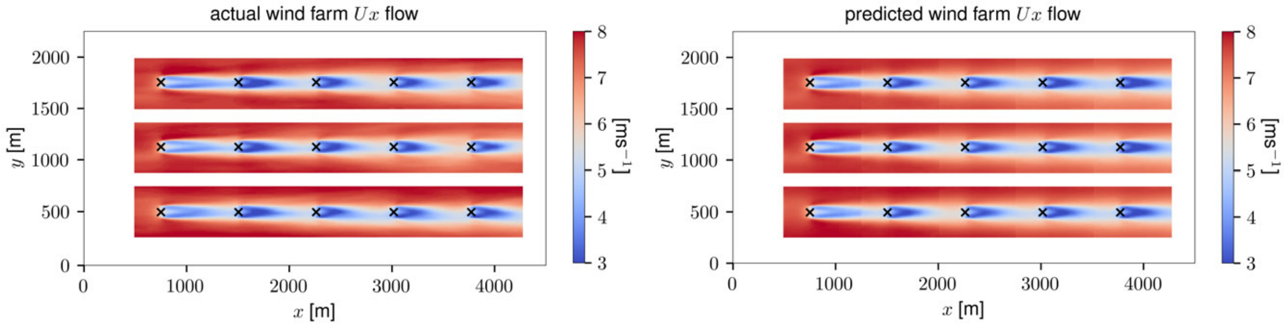


Figure 1: Mean streamwise velocity component at hub height. LES-ALM data (left) and GAN prediction (right). Time window: 4000[PK1.1]. Precursor simulation: 7.7 m/s. Inflow angle: 0°.

in a rectangular region around that turbine. The GAN therefore learns the mapping from local inflow to the surrounding wake, including wake recovery and lateral spreading, for a range of inflow conditions and turbine positions in the farm.

The surrogate follows a conditional GAN architecture. The generator receives the upstream velocity profile and produces a two-dimensional map of the mean hub-height wind speed in a region that extends several rotor diameters upstream and downstream of a turbine. The discriminator processes paired inflow–wake fields and learns to distinguish synthetic wakes produced by the generator from reference wakes extracted from the LES-ALM simulations. Training minimises a combined loss that balances an adversarial term, which encourages realistic global wake structures, and a mean-square error term, which penalises local deviations from the LES-ALM fields.

The experimental analysis explores several training setups by varying the averaging time window used to construct the mean fields and the weight of the error term in the loss [1]. Performance is evaluated using both image-based measures and flow-specific quantities such as wake-centre position and velocity deficit. The best configuration achieves low mean errors when predicting the wake around an individual turbine given the local inflow.

Figure 1 illustrates a representative comparison between the LES-ALM reference and the GAN prediction for the mean streamwise velocity at hub height in a 15-turbine farm. The surrogate reproduces the location and depth of the velocity deficit with moderate errors over most of the wake region, including the gradual recovery of wind speed downstream. Differences concentrate near the rotor plane and at the edges of the wakes, where gradients are sharp and small misalignments translate into visible discrepancies.

The study also examines whether the surrogate can reconstruct the hub-height velocity field across the entire farm using only the inlet profile. In that test, the model is applied sequentially along each row, using the predicted profile at one turbine as input for the next. Errors accumulate downstream and become significant in the last rows, especially for shorter averaging windows where inflow variability plays a stronger role. The results indicate that purely data-driven GAN surrogates capture local wake features reliably, but global consistency across multiple wake interactions remains challenging.

Despite these limitations, the surrogate reduces evaluation time from days to seconds once trained. That speed enables integration into optimisation and control studies that require thousands of flow evaluations, such as layout optimisation or yaw control strategies. In the ongoing project, such data-driven models serve as a reference point for the development of PI-GANs that will incorporate momentum conservation and turbulence closures directly into the training objective, with the goal of improving robustness and generalisation.

The combination of PI-GAN concepts with high-fidelity LES-ALM data and evolutionary optimisation defines a research direction in which generative models not only reproduce existing simulations [2], but also explore new operating regimes under explicit physical constraints. The wind-turbine wake surrogate in [1] demonstrates the feasibility of GAN-based flow prediction in wind farms and motivates further work on physics-informed variants that can better handle complex wake interactions in realistic conditions.

Link:
[L1] https://jamaltoutouh.github.io/pinns/

References:
[1] M. Bove, et al., "Estimation of wind turbine wakes with generative-adversarial networks," Journal of Physics: Conference Series, 2505, 012053, 2023.
[2] L. Yang, D. Zhang, G.E. Karniadakis: "Physics-informed generative adversarial networks for stochastic differential equations," SIAM J. Sci. Comput. 42(1), A292–A317, 2020.

Please contact:
Jamal Toutouh, Universidad de Málaga, Spain
jamal@uma.es

Bridging Modelling Scales with AI: Deep Learning – Enhanced Multi-Scale Simulations of Molecular Systems

by Eleftherios Christofi (The Cyprus Institute) Vagelis Harmandaris (The Cyprus Institute, University of Crete and FORTH-IACM)

By enhancing multi-scale molecular simulations with deep learning, AI is enabling researchers to bridge modelling scales that were once out of reach. This synergy opens new possibilities for understanding complex materials and accelerating scientific discovery.

Understanding complex macromolecular systems requires computational tools that can capture their behaviour on a wide range of spatio-temporal (length and time) scales. Classical atomistic simulations provide a valuable tool for exploring the structure and dynamic evolution of polymeric systems at the molecular scale, but they face fundamental limitations on the accessible length and time scales, mainly due to the long relaxation times required for equilibration. To overcome such limitations, a number of different approaches have been employed to model systems across different levels of detail. Hierarchical particle-based computational modelling of macromolecular materials across various scales is based on the systematic exchange of information between different (e.g. microscopic and mesoscopic) representations of atoms and molecules to capture material properties for a wide range of spatial and temporal scales. By considering models at different scales simultaneously, one hopes to develop an approach that shares the efficiency of the coarser models, as well as the accuracy of the microscopic (finer) ones.

Therefore, such approaches are nowadays applied to a wide range of polymer-based molecular systems in various fields such as physics, chemistry, biology, materials science, and engineering. A common characteristic of all the above systems is that they are characterised by a very broad range of spatio-temporal scales. Hence, to predict their properties and obtain a fundamental understanding of their behaviour, simulation methods over different scales are necessary.

Artificial intelligence (AI), and deep learning in particular, is now offering new opportunities to advance these multi-scale approaches. Neural networks can learn structural patterns directly from simulation data, allowing them to represent complex molecular information in ways that traditional rules or functional forms cannot easily achieve. This makes AI a valuable complement to classical modelling, especially in situations where information is lost during coarse-graining or when atomistic simulations become too costly.

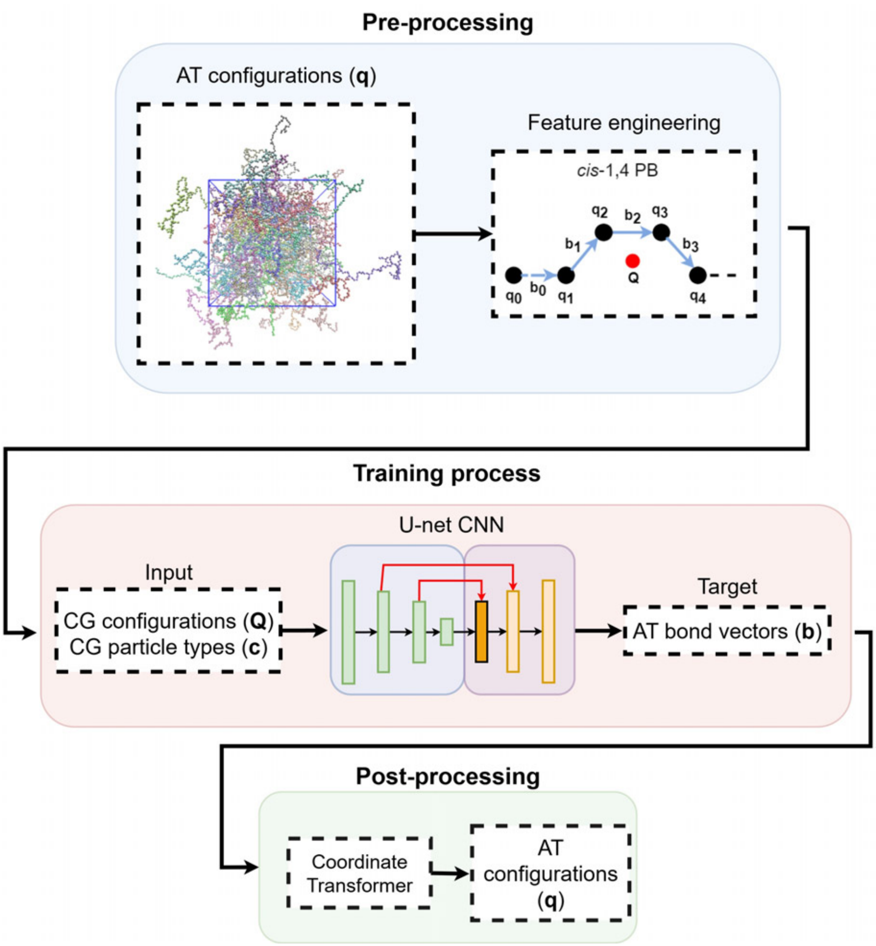


Figure 1: Schematic representation of the systematic back-mapping of the coarse-grained multi-component molecular system via deep learning models.

An important direction involves moving upwards in resolution. To expand the range of accessible scales via atomistic simulations, coarse-grained (CG) models that reduce the dimensionality of the physical system under study have been developed. In systematic bottom-up CG models, groups of atoms are lumped together into particles that are typically denoted as "superatoms" or beads. Currently, such approaches are applied to a multitude of molecular systems, such as proteins and synthetic polymers. However, when a molecular system is simplified to a coarser representation, much of its atomistic detail is removed. Therefore, in order to obtain information on properties that depend on the microscopic structure, an atomic-level resolution is necessary. Restoring this missing information is a difficult task, yet it is essential when fine-scale accuracy is needed. Recently, we have developed a general DL-based approach for back-mapping (re-introducing atomic detail) in CG multi-component macromolecular systems, based on convolutional neural networks (CNN) that are trained directly on atomistic descriptors (see Figure 1). The new DL model is capable of creating representative well-equilibrated atomistic configurations in three-dimensional Cartesian space of multi-component polymeric materials of high molecular weight [1, L1] (see Figure 2). To improve the accuracy of the deep learning model, the loss function of the neural network is augmented with several penalizing terms based on prior knowledge of the physical properties of the molecular systems under study, such as bond lengths, bond angles, and dihedral angles. In addition, the proposed approach avoids tedious and labor-intensive bookkeeping of molecular details during the reconstruction by separating configurational information from molecular topology and force fields.

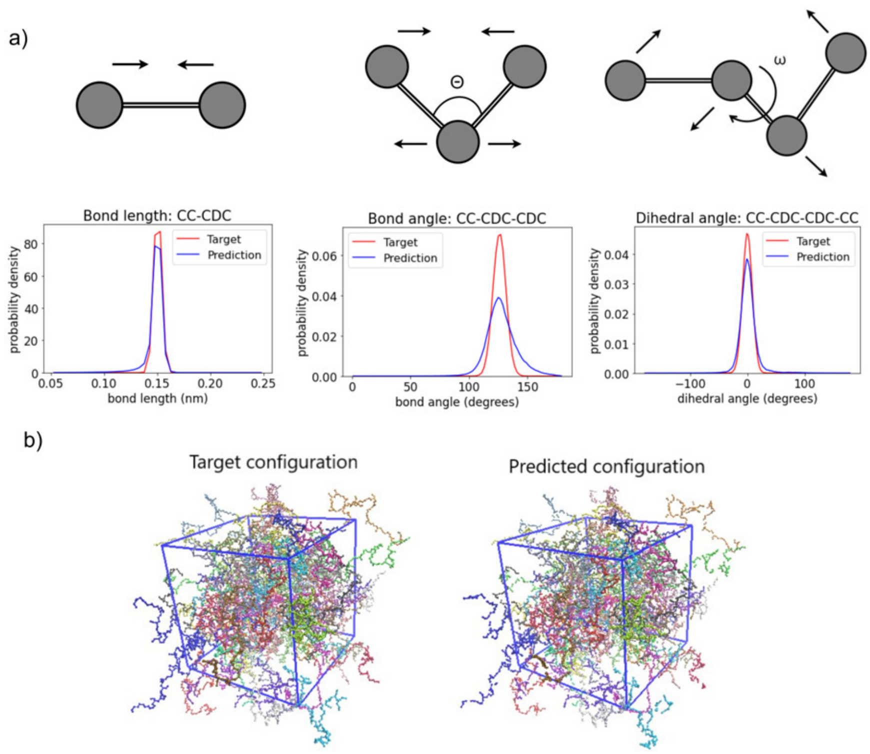


Figure 2 (a) Comparisons of bond lengths, bond angles, and dihedral angles among target atomistic configurations and predictions obtained via deep learning models. (b) Target and predicted (through deep learning models) snapshots of a bulk polymeric system generated through deep learning models.

Multi-scale modelling can also be used to make predictions that move downwards in resolution, from detailed molecular structure towards macroscopic behaviour. This is particularly relevant for polymer nanocomposites, where local structural heterogeneities and polymer–filler interactions strongly influence the overall mechanical response. Atomistic simulations are capable of resolving per-atom stress and strain fields, yet the computational cost of generating these fields for multiple systems and load cases is substantial. To address this challenge, a data-driven model was developed that predicts these mechanical fields directly from structural descriptors extracted from atomistic trajectories [2, L2]. The model learns to reproduce the spatially varying mechanical response across the matrix, interphase, and nanoparticle regions, allowing researchers to examine how filler content, polymer configuration, and interfacial structure affect the emergent mechanical properties without performing new large-scale simulations for each case.

Together, these directions demonstrate how AI can support multi-scale modelling by improving the flow of information between different levels of resolution. Deep learning assists in reconstructing the fine-scale atomistic structure when needed and accelerates the prediction of macroscopic properties originating from microscopic features. In this way, AI driven models help balance efficiency and accuracy, allowing researchers to explore large systems while still retaining access to detailed molecular information. Building on these methodologies, future work will focus on developing models to explore the chemical design space, which can be used to support the in-silico design of materials with targeted macroscopic properties.

As AI-powered modelling continues to mature, its integration with multi-scale approaches is expected to deepen, opening new possibilities for the design and understanding of advanced molecular materials. The combination of hierarchical modelling and machine learning provides an effective route towards bridging previously inaccessible scales and offers a promising foundation for future advances in macromolecular simulation.

Our research team is an interdisciplinary team composed of chemical engineers and applied mathematicians. We have long-standing experience in model reduction methods, mathematical and computational modelling of complex systems, and scientific machine learning.

Three different institutes are involved in the research project: the University of Crete (Greece), the Institute of Applied and Computational Mathematics – FORTH (Greece) and the Cyprus Institute (Cyprus). We also have a long-term' active collaboration with research institutes around the world.

Links:
[L1] https://github.com/SimEA-ERA/CNN-BackMap-CG
[L2] https://github.com/SimEA-ERA/DeepNANOMEC

References:
[1] E. Christofi et al., "Deep convolutional neural networks for generating atomistic configurations of multi-component macromolecules from coarse-grained models", J. Chem. Phys., 2022, 157, 184903. https://doi.org/10.1063/5.0110322
[2] E. Christofi et al., "Physics-informed deep learning model for predicting the heterogeneous mechanical properties of polymeric nanostructured materials", SSRN, 2025. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5372731

Please contact:
Vagelis Harmandaris, Eleftherios Christofi
The Cyprus Institute, Cyprus, University of Crete, Greece, FORTH-IACM, Greece
v.harmandaris@cyi.ac.cy, e.christofi@cyi.ac.cy

SmartCityHub, an Enabler for Trusted AI Solutions

by Andrés Meléndez Imaz, Thomas Tamisier (LIST)

SmartCityHub acts as a structured bridge between cities and AI solutions providers through a multi-phase lifecycle designed to minimise risk and maximize scalability. It provides a testing platform within a controlled environment and ensures regulatory compliance and safety before deployment.

Smart Cities approaches define urban innovation strategies and implementation guidelines relying on digitalisation, coupled notably with the latest artificial intelligence (AI), interoperability and Local Digital Twin (LDT) technologies. To translate these strategies into practical and impactful components, advanced technologies must be validated with regard to functional and regulatory requirements. The EU has launched major investments to accelerate the development of responsible AI in Europe, in particular in the domain of Smart Cities and Communities [1], with the Testing and Experimenting Facilities (TEF) programme, which offers support to innovators and communities of users to test latest AI-based software and hardware technologies at scale in real-world environments. As part of this broad European initiative and within the project CitCom.ai [L1], LIST has established SmartCityHub (SCH) as a reference platform to conduct pilot experiments and link cities and AI innovators to create scalable AI-powered services [L2]. SCH enables on the one hand cities to explore, validate, and pilot smart city technologies. On the other hand, by linking municipalities in Luxembourg with artificial intelligence innovators across Europe, it supports collaborative experimentation of AI-driven solutions under transparent and trustworthy conditions.

More specifically, SmartCityHub helps to define the scope for innovation in cities and communities, as well as to identify opportunities for new services. The process starts by analysing functional needs, digital maturity and feasibility, as regards technical and practical deployment. A tailored set of activities is then proposed according to the idea's maturity and customer requirements, including:
- The Digital Opportunity Assessment, which identifies digitalisation opportunities for companies (use cases) and build the best suited approach to test them. The outputs are requirements for a proof-of-concept.
- The co-creation workshops, customised for communities and preceded by research on local challenges and strategic milestones, to help prioritise use cases for data analytics, AI, and Local Digital Twins.
- The technological proof-of-concept phase, which enables organisations to test solutions before investing, through the development of a tailored proof-of-concept to explore the potential of data analytics and AI.

In a subsequent phase, SCH, as part of LIST's role as a Research and Technology Organisation (RTO), acts as a neutral and trustworthy partner to refine concrete use cases (primarily, though not exclusively, derived from workshop outcomes) and set-up a series of experiments onboarding cities and AI solution providers. We first measure and analyse the technological gaps (notably in terms of data provisioning, interoperability), followed by the deployment of cutting-edge tools; we then guide the executing of experiments in a controlled environment with relevant data, before analysing the results to extract valuable insights from the adopted solution. Additionally, matchmaking services connect cities with AI innovators based on well-defined use cases [2].

Finally, two illustrative examples demonstrate how the SCH collaborates with cities to implement and validate innovative technologies.

In the first collaboration, a city deployed an AI-based chatbot to assist citizens by answering questions regarding administrative procedures. In close interaction with the AI provider of the solution, SCH conducted a technical evaluation based on LIST's AI Sandbox [L3], to assess and benchmark the chatbot's large language models with respect to bias, fairness, ethical risks and to ensure their safe adoption by the public. The evaluation was done through a modular architecture that served as a communication interface between the test suite and the AI-based chatbot. This interface transmitted structured test prompts to the target AI system and collected the corresponding outputs (such as response texts and quantitative evaluation metrics) for subsequent analysis.

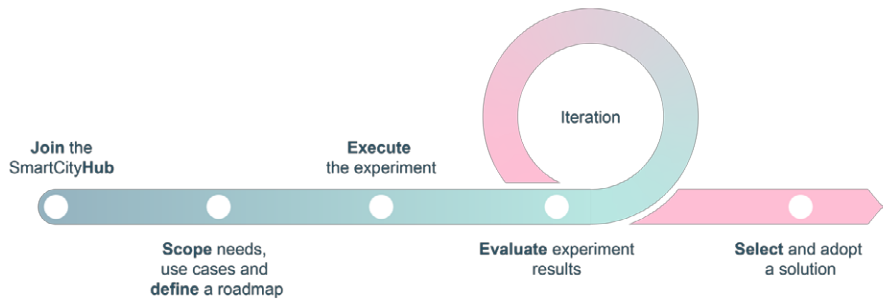


Figure 1: Lifecycle of SmartCityHub interaction.

In the second collaboration, the city of Differdange (an industrial municipality aiming to achieve carbon neutrality by 2030) defined concrete use cases focused on energy efficiency through a needs-assessment workshop, SCH developed a proof-of-concept based on a Local Digital Twin that included:
- A dashboard aggregating data on public buildings (such as location, size, structural propertiesand historical electricity consumption).
- Simulation and what-if scenarios to identify optimal locations for electric vehicle charging stations and to estimate the potential energy production of rooftop solar panels.

These collaborations showed how SmartCityHub operationally supports different stakeholders in delivering concrete AI driven Smart City solutions.. As a European Test and

Experimentation Facility, SCH offers a full range of services to connect city demand with latest AI innovations, ensuring trusted adoption and replication across different contexts through tested and tailored solutions.

Links:
[L1] https://citcomtef.eu
[L2] https://smartcityhub.list.lu
[L3] https://ai-sandbox.list.lu

References:
[1] Smart Cities: A Key to a Progressive Europe, Foundation for European Progressive Studies (FEPS), Brussels, Belgium, Mar. 2020. [Online]. Available: https://feps-europe.eu/wp-content/uploads/2020/03/Smart-Cities-A-Key-to-a-Progressive-Europe.pdf
[2] Les communes luxembourgeoises face à la transition numérique et écologique, Infogreen Luxembourg. [Online]. Available: https://www.infogreen.lu/les-communes-luxembourgeoises-face-a-la-transition-numerique-et-ecologique.html

Please contact:
Andrés Meléndez Imaz, LIST, Luxembourg
andres.melendez@list.lu

Thomas Tamisier, LIST, Luxembourg.
thomas.tamisier@list.lu

Analyzing Social Engineering Automation

by Dominik Dana (St. Pölten University of Applied Sciences), Sebastian Schrittwieser (University of Vienna, JRC AsTra), Peter Kieseberg (St. Pölten University of Applied Sciences)

In order to analyse the practical capabilities and limitations of automated social engineering, we conducted a practical study covering more than 140 open-source tools. In order to achieve comparability, we provided an abstract model based on extracting the relevant aspects of the most important attack modelling frameworks.

In recent years, the development of social engineering (SE) techniques has been accompanied by the emergence of a wide range of automated tools. The integration of artificial intelligence (AI) into these tools has further enhanced their capabilities, particularly by reducing the manual effort required to conduct targeted attacks, as attackers can now design personalized attack paths based on information collected about specific individuals. Despite this progress, most available tools remain specialized, addressing only selected tasks rather than the entire attack chain. Furthermore, several different approaches for modelling social engineering have been proposed in the literature, each with its specific advantages; however, this reduces comparability of said tools.

To structure the analysis, a technical social engineering model was developed by synthesizing the commonalities of established frameworks. The study aims to illustrate the degree to which automation in SE attacks is feasible, while also assessing the functionalities and reliability of the tools under investigation.

Automation in social engineering evolves rapidly, while academic approaches and real-world implementations often diverge substantially. Consequently, a purely literature-based approach would not have provided sufficient insight, thus we followed a more practical approach by evaluating 140 different open-source tools within a practical application setting, illustrating the degree to which automation in SE attacks is feasible, and assessing the functionalities and reliability of the tools[1]. More specifically, our research addresses the following research questions (RQs):
- RQ1: To what extent are freely available SE tools already automated, and what are the implications for social engineering?
- RQ2: Which phases of social engineering can be supported by existing tools?
- RQ3: How do different tools interact, and are there tool suites that support the entire SE process?
- RQ4: How reliable are the results produced by these tools?

Figure 1: Generalized SE Model – Overview [1].

To answer these questions, a generalized model for social engineering attacks was derived (Figure 1). This model synthesizes the most prominent SE frameworks, including the Cyber Kill Chain, the Social Engineering Cycle, the Social Engineering Lifecycle, the Social Engineering Pyramid, the Social Engineering Attack Framework, the Cycle of Deception, the Social Engineering Attack Spiral, the Session and Dialogue-Based Framework, and the Phase- and Source-Based Model. The generalized model enables consistent mapping and comparison of tool functionalities across heterogeneous frameworks.

The study showed that freely available SE tools exhibit automation primarily in recurring queries and search operations. Automation capabilities are enhanced when tools rely on application programming interfaces (APIs), as outputs can be further processed automatically. A fully automated end-to-end solution, however, was not identified. Automation is also prevalent in attack execution and preparation, where tools are designed for ease of use. The majority of automation was observed in information retrieval technologies, likely influenced by the strong open-source intelligence (OSINT) community. The analyzed frameworks vary with respect to the number and definition of phases, but overall, reconnaissance and attack-execution phases benefit most from automation. Due to structural differences between frameworks, direct mapping of tools to all models was not feasible. For this reason, the generalized model served as the reference framework for tool classification.

Even though the level of automation is very good in some parts of the SE lifecycle, especially considering reconnaissance, manual intervention remains necessary for record selection, validation, and formatting before data can be processed by subsequent tools. Special tool suites such as Maltego [L1] and Lampyre [L2], which provide extensibility through plug-ins, cover a large portion of the SE process. While they do not offer complete end-to-end automation, they provide stable and reliable support across multiple phases. Tool reliability often depends strongly on their operational mode. Some tools rely on archived databases or previously crawled and scanned websites, whereas others query live data sources. These differences affect both accuracy and timeliness of results.

The analysis demonstrates that automation is most advanced in the information-gathering phase. This is reflected in the large number of tools available, the rapid release cycle of new applications and updates, and the linguistic diversity of implementations. The study revealed that information retrieval within the European Union has become more challenging since the introduction of the General Data Protection Regulation (GDPR) [2], while many web-based retrieval applications continue to provide results predominantly for the United States.

Most free-to-use tools impose query limits, restricting large-scale automation, and furthermore, the integration of AI remains relatively immature and falls short of expectations. Current tools thus automate individual tasks effectively but lack seamless integration into a fully automated SE pipeline, but future developments in API access, data integration, and AI-driven contextual analysis may eventually enable end-to-end automation of SE attacks, raising critical implications for both attackers and defenders.

For future work, we will expand these experiments and plan to continuously evaluate new developments in the field. With the advent of LLMs for low budgets, especially the automation of phishing and vishing comes within the reach of non-experts, which will likely change the attacker profiles in these areas.

Links:
[L1] https://www.maltego.com/
[L2] https://lampyre.io/

References:
[1] D. Dana, S. Schrittwieser, P. Kieseberg, “Automated Social Engineering Tools-Overview and Comparison with Respect to Capabilities and Detectability”, ICCGI 2024.
[2] European Parliament and Council, “Regulation (EU) 2016/679 (General Data Protection Regulation),” Official Journal of the European Union, 2016.

Please contact:
Peter Kieseberg
St. Pölten University of Applied Sciences, Austria
peter.kieseberg@fhstp.ac.at

CALL FOR PARTICIPATION

2nd AIOTI Workshop on Semantic Interoperability for Digital Twins

Frankfurt (Oder), Germany, 24–25 March 2026

The 2nd AIOTI Annual Workshop on Semantic Interoperability for Digital Twins, with a special focus on reusability and compliance, will take place on 24–25 March 2026 in Frankfurt (Oder), Germany, hosted by IHP GmbH – Leibniz-Institut für innovative Mikroelektronik.

Organised under the umbrella of AIOTI, the workshop will bring together representatives from industry, standards bodies, and research to discuss practical approaches to achieving semantic interoperability for Digital Twins. A key outcome of the event will be the development of a Manifesto for Reusable Ontologies, aimed at supporting scalable, vendor-neutral, and compliant Digital Twin ecosystems.

Topics include semantic interoperability for Digital Twins; reusable and FAIR ontologies; compliance, validation, and conformance testing; governance and lifecycle management of semantic assets; cross-domain interoperability; standards and vendor neutrality; and the role of AI, including large language models and agentic AI, in semantic technologies.

Expressions of interest to participate, as well as proposals for talks, posters, demonstrations, panels, breakout sessions, and lightning talks, are invited.

The submission deadline is 13 February 2026. If an extension of the deadline is required, please contact Martin Bauer.

Links:
https://aioti.eu/event/2nd-aioti-workshop-on-semantic-interoperability-for-digital-twins/
Submission form: https://kwz.me/hIh

Please contact:
Call and programme:
Martin Bauer, martin.bauer@neclab.eu

Venue and logistics:
Krzysztof Piotrowski
piotrowski@ihp-microelectronics.com

CALL FOR WORKSHOP PROPOSALS

SAFECOMP 2026 45th International Conference on Computer Safety, Reliability and Security

Valencia, Spain, 22–25 September 2026

The 45th edition of the International Conference on Computer Safety, Reliability and Security (SafeComp 2026), to be held in Valencia in September 2026, will focus on the theme "Engineering safe and sustainable computing systems", addressing the challenge of combining safety, security and sustainability in computing infrastructures. Founded in 1979, this conference brings together experts to share advances, experiences, and trends in the safety, reliability, and sustainability of critical systems.

SafeComp 2026 invites proposals for co-located workshops that will complement the main symposium by offering forums for focused discussions on emerging and specialised topics related to the safety, reliability, and sustainability of critical systems. Workshops may serve as platforms for early-stage idea exchange, in-depth research discussions, work-in-progress reports from research projects, or collaborative working sessions to address current challenges in the field.

Workshops are expected to attract both academic and industrial participants, including researchers, practitioners, and students. Accepted full workshop papers (6–12 pages), reviewed according to Springer LNCS guidelines by at least three independent reviewers, will be published in the SafeComp 2026 Workshops Proceedings. All workshops will be held in parallel on 22 September, ahead of the main conference. Authors of workshop papers are kindly invited to participate in the SafeComp main conference.

Submission Guidelines for Workshops Proposals

Prospective workshop organizers should submit a 2–3 - page proposal including:
- Workshop Title
- Description: goals, scope, relevance to SafeComp, and intended audience.
- Workshop Format: e.g., full/short paper presentations, posters, invited talks, panels, demos, hands-on sessions, etc.
- Submission and Review Process: details on paper length, review criteria, and promotional strategies (following general Springer LNCS publishing rules).
- Duration: half-day or full day.
- History: information about previous editions (if applicable).
- Workshop Committee: short bios of the main workshop chairs, previous experience in similar roles, and a tentative list of Workshop Program Committee members.

Important dates

- Workshop Proposal Submission Deadline: 23 February 2026
- Notification of Acceptance: 20 March 2026

Proposal submission

Proposals should be submitted via email to safecomp2026-workshops@upv.es and erwin.schoitsch@ait.ac.at.

Link:
https://safecomp2026.webs.upv.es/

Please contact:
Erwin Schoitsch, AIT, Austria
General SAFECOMP 2026 Workshop Chair
erwin.schoitsch@ait.ac.at

CALL FOR PAPERS

SAFECOMP 2026

Valencia, Spain,22-25 September 22-25, at the Universitat Politècnica de València!

SafeComp is an annual international conference addressing state-of-the-art research, industrial experience, and emerging trends in the safety, security, and reliability of critical computer applications. The 2026 theme, "Engineering safe and sustainable computing systems", reflects the growing need to combine safety assurance with sustainability in computing infrastructures.

Established in 1979 by EWICS TC7, SafeComp has long contributed to advances in dependable and safety-critical systems. As a single-track conference without parallel sessions, it offers a strong forum for exchange and networking on emerging methods and practical solutions.

There are three opportunities for papers:
- Regular papers: Cover all aspects of the development, assessment, operation, and maintenance of safety-critical computer systems.
- Position papers: Present viewpoints and perspectives on future directions, rather than reporting completed research.
- Workshop papers: Submitted to workshops selected through a Call for Workshop Proposals (see examples from last year

Important dates (Deadlines)

- Workshop proposals: 23 February 2026
- Full paper submission: 27 February 2026
- Author notification: 24 April 2026
- Camera-ready submission: 7 June 2026
- Workshops: 22 September 2026
- Main conference: 23–25 September 2026

Links:
https://safecomp2026.webs.upv.es/
https://kwz.me/hGz (Leaflet)

PRE-ANNOUNCEMENT

21st International EWICS / ERCIM Workshop on Dependable Smart Embedded Cyber-Physical Systems and Systems-of-Systems (DECSoS 2026)

Valencia, Spain, 22–25 September 2026

The DECSoS Workshop continues an initiative of the ERCIM-DES Working Group and has been successfully sustained over the years. Held in conjunction with SAFECOMP 2026, it follows a well-established tradition dating back to 2006. Originally focused on embedded systems and their dependability properties, the workshop evolved to emphasise cyber-physical systems, reflecting closer links to physics, mechatronics, and interaction with partly unpredictable environments.

A current focus is the trustworthiness of smart and autonomous systems composed of cognitive CPS integrated into IoT infrastructures. Key issues include digitalisation, in particular extensive V&V activities on Digital Twins, and new paradigms in software and systems engineering, such as functional safety in systems including AI-based elements.

Given the considerable societal impact of these technologies, dependability must be addressed holistically, encompassing resilience, sustainability, and ethically aligned design. Cognitive systems, CPS, and IoT are priority research areas in Horizon Europe and in public–private partnerships such as the Chips Joint Undertaking.

Sessions

Sessions are planned on:
- Dependable and resilient embedded systems, systems of cyber-physical systems
- Highly automated (autonomous) Systems and Robotics
- AI and autonomy: functional safety, cybersecurity, and human–machine teaming
- Medical devices and healthcare: safety, security, and conformity assessment
- Smart Anything Everywhere and the Internet of Things (IoT)
- Digitalisation towards Society 5.0 (Industry 5.0, Farming 5.0, smart mobility, digital cities, smart health), with particular attention to environmental, sovereignty, sustainability, security, human aspects, and ethically aligned design.

The sessions are addressing thematic topics such as:
- Multi-core platforms and mixed criticality systems
- Safety and security co-engineering for trustworthiness
- Validation and Verification, multi-concern and modular assurance
- Domain-specific critical applications (industrial, mobility, medical devices and others)
- Standardization (interoperability, trustworthiness), certification and ethical concerns.

The topics cover aspects ranging from concepts to deployment and maintenance. To remain distinct from the SAFECOMP conference mainstream, the workshop focuses on reports of ongoing work in progress, fostering discussion, experience exchange, and the presentation of unconventional subtopics. Reports on European or national research projects, as part of required dissemination, as well as industrial experience reports, are particularly welcome.

All papers will be reviewed by at least three reviewers. The workshop proceedings will be published as a complementary volume to the SAFECOMP Proceedings in Springer LNCS. Papers (6–12 pages) should be prepared according to the Springer LNCS formatting guidelines. Submission will be via EasyChair (link to be announced in February 2026):

Deadlines:
- Full paper submission: 4 May 2026
- Notification of acceptance: 18 May 2026
- Camera-ready submission: 7 June 2026
- Workshop: 22 September 2026.

The Programme Committee is composed of EWICS and ERCIM members.

Please contact the Workshop and Programme Committee Chairs:

Erwin Schoitsch
AIT Austrian Institute of Technology
erwin.schoitsch@ait.ac.at

Amund Skavhaug, NTNU, Norway
amund.skavhaug@ntnu.no

POSITION PAPER

Reclaiming Software Engineering as a Key Enabler of the Digital Age

A joint Informatics Europe / ERCIM Working Group on Software Research has published the position paper “Reclaiming Software Engineering as the Enabling Technology for the Digital Age.” The paper argues that software engineering must be recognised as a foundational enabling technology underpinning advances across artificial intelligence, data-driven systems, cyber-physical systems, and digital infrastructures.

The authors highlight that, despite its central role, software engineering is often treated as secondary to other digital technologies. The paper calls for renewed investment in software research addressing scalability, dependability, sustainability, and human-centric development, and for stronger integration of software engineering perspectives into European research and innovation strategies. Repositioning software engineering, the paper argues, is essential for ensuring that future digital systems are trustworthy, maintainable, and aligned with societal needs.

Reference:

T. E. Vos, et al., “Reclaiming Software Engineering as the Enabling Technology for the Digital Age,” arXiv preprint, arXiv:2601.14861, 2026.
https://arxiv.org/pdf/2601.14861

Call for Proposals

Dagstuhl Seminars and Perspectives Workshops

Schloss Dagstuhl – Leibniz-Zentrum für Informatik is accepting proposals for scientific seminars/workshops in all areas of computer science, in particular also in connection with other fields.

If accepted, the event will be hosted in the seclusion of Dagstuhl's well known, own, dedicated facilities in Wadern on the western fringe of Germany. Moreover, the Dagstuhl office will assume most of the organisational/ administrative work, and the Dagstuhl scientific staff will support the organizers in preparing, running, and documenting the event. Thanks to subsidies the costs are very low for participants.

Dagstuhl events are typically proposed by a group of three to four outstanding researchers of different affiliations. This organizer team should represent a range of research communities and reflect Dagstuhl's international orientation. More information, in particular details about event form and setup, as well as the proposal form and the proposing process, can be found on

https://www.dagstuhl.de/dsproposal

Schloss Dagstuhl – Leibniz-Zentrum für Informatik is funded by the German federal and state government. It pursues a mission of furthering world class research in computer science by facilitating communication and interaction between researchers.

Important Dates

- *Next submission period: April 1 to April 15, 2026*
- *Seminar dates: April 2027 and March 2028 (tentative).*

CALL FOR PAPERS

FMICS 2026: 31st International Conference on Formal Methods for Industrial Critical Systems

Liverpool, UK, 2-4 September 2026

FMICS is the annual conference of the ERCIM Working Group on Formal Methods for Industrial Critical Systems, and it is the key conference in the intersection of industrial applications and Formal Methods.

The aim of the FMICS conference series is to provide a forum for researchers and practitioners who are interested in the development and application of formal methods in industry. FMICS brings together scientists and engineers who are active in the area of formal methods and interested in exchanging their experiences in the industrial usage of these methods. The FMICS conference series also strives to promote research and development for the improvement of formal methods and tools for industrial applications.

As in previous years, FMICS is part of the CONFEST umbrella umbrella event consisting of three main conferences CONCUR, FMICS and QEST+FORMATS, and affiliated workshops. The 2026 edition will take place on the main campus of the University of Liverpool, on September 1-5, 2026.

Topics of Interest

- Formal specification, including specification elicitation, validation, debugging, sanity checking, revision, coverage, and explainability.
- Case studies and experience reports on industrial applications of formal methods, focusing on lessons learned or identification of new research directions.
- Methods, techniques, and tools to support automated analysis, certification, debugging, learning, optimization, and transformation of complex, distributed, real-time, embedded, mobile, and autonomous systems.
- Verification and validation methods (model checking, theorem proving, SAT/SMT constraint solving, abstract interpretation, etc.) that address shortcomings of existing methods with respect to their industrial applicability (e.g., scalability and usability issues, tool qualification, and certification).
- Transfer to industry and impact of adoption of formal methods on the development process and associated costs in industry as well as application of formal methods in standardization and industrial forums.

Deadlines

The paper submission deadline is 17 April 2026. Accepted papers will be included in the Conference Proceedings published as part of Springer's Lecture Notes in Computer Science series.

As in best FMICS tradition, the paper with the best contributions to Software Science and Technology will be honoured with the EASST best paper award.

Invited Speakers

- Julia Badger (NASA Johnson Space Center, USA)
- Colin O'Halloran (D-RisQ, UK)

PC Chairs

- Peter Gorm Larsen (Aarhus University, Denmark)
- Kristin Yvonne Rozier (Iowa State University, USA)

More information:
https://confest-2026.github.io/fmics/

Submission link:
https://easychair.org/conferences?conf=fmics2026

CALL FOR NOMINATIONS

Fraunhofer–Bessel Research Award

Since 2005, the Alexander von Humboldt Foundation and the Fraunhofer-Gesellschaft have jointly awarded the Fraunhofer–Bessel Research Award to internationally recognised researchers for outstanding scientific achievements. The award is endowed with €45,000 and enables recipients to carry out collaborative research at a Fraunhofer Institute in Germany for a total period of six to twelve months, which may be divided into multiple stays.

Institute directors or senior researchers at Fraunhofer Germany may nominate leading scientists who completed their doctorate less than 18 years ago; self-nominations are not permitted. The programme is open to researchers of all nationalities who have been predominantly living and working outside Germany for at least five years at the time of nomination. Nominations can be submitted throughout the year. Researchers interested in being nominated are invited to contact the Fraunhofer Institute they wish to collaborate with, or alternatively Lisa Göbl or Annika Wust for further guidance.

More information:
https://www.humboldt-foundation.de/en/apply/sponsorship-programmes/fraunhofer-bessel-research-award

Please contact:
Lisa Göbl or Annika Wust,
Fraunhofer-Gesellschaft, Germany
lisa.goebl@zv.fraunhofer.de
annika.wust@zv.fraunhofer.de

In Memoriam: Keith G. Jeffery

Keith G. Jeffery passed away on 15 January 2026. He served the European research community, and ERCIM in particular, with long-standing commitment over several decades.

Keith took office as ERCIM President on 1 January 2005, at a time when he was Director of IT at CCLRC, the UK member organisation of ERCIM (later STFC). He firmly believed that only by working together in a coordinated manner, ERCIM could play a meaningful role in the European knowledge society. This conviction guided his presidency until its conclusion at the end of 2013.

During his presidency, ERCIM grew into a widely recognised key player in European ICT research. Keith steered the Consortium through a number of important structural changes. In particular, he supported the decision to open ERCIM membership to allow multiple members per country, enabling excellent ICT research institutions to participate directly in ERCIM activities rather than via one national intermediary. This step significantly broadened ERCIM's scientific base and strengthened its role at the European level.

Keith's academic background was in geology, in which he earned his PhD. His research led him at an early stage to the field of information systems, shaping a career that focused on database systems, scientific computing, research information systems, and e-infrastructures.

Before and during his leadership roles in ERCIM, Keith gained extensive experience within European and national research infrastructures. At the Rutherford Appleton Laboratory, he led research groups and divisions working on database systems, knowledge and software engineering, and large-scale research services. As Director of IT at CCLRC (1998–2006) and later at STFC (2006–2013), he was responsible for major research IT operations and international strategy, while remaining actively involved in European ICT research projects and expert groups.

Keith gained experience across all major governance structures of ERCIM. He chaired the Database Research Working Group from 1991 to 1999, served as a member of the Executive Committee from 1994 to 1998, and joined the Board of Directors in 1998, where he remained active for many years.

After stepping down as ERCIM President, Keith continued to support ERCIM activities. Following his retirement, he served as a scientific coordinator in several EU-funded projects, contributing his experience in coordination, knowledge transfer, and European research collaboration.

Keith G. Jeffery will be remembered with great appreciation for his leadership, professionalism, and his sustained commitment to strengthening European research through cooperation.

ERCIM – the European Research Consortium for Informatics and Mathematics is an organisation dedicated to the advancement of European research and development in information technology and applied mathematics. Its member institutions aim to foster collaborative work within the European research community and to increase co-operation with European industry.

ERCIM is the European Partner of the World Wide Web Consortium.

Consiglio Nazionale delle Ricerche
Area della Ricerca CNR di Pisa
Via G. Moruzzi 1, 56124 Pisa, Italy
www.iit.cnr.it

I.S.I. – Industrial Systems Institute
Patras Science Park building
Platani, Patras, Greece, GR-26504
www.isi.gr

Centrum Wiskunde & Informatica
Science Park 123,
NL-1098 XG Amsterdam, The Netherlands
www.cwi.nl

Norwegian University of Science and Technology
Faculty of Information Technology, Mathematics and Electrical Engineering, N 7491 Trondheim, Norway
http://www.ntnu.no/

Fonds National de la Recherche
6, rue Antoine de Saint-Exupéry, B.P. 1777
L-1017 Luxembourg-Kirchberg
www.fnr.lu

RISE SICS
Box 1263,
SE-164 29 Kista, Sweden
http://www.sics.se/

Foundation for Research and Technology – Hellas
Institute of Computer Science
P.O. Box 1385, GR-71110 Heraklion, Crete, Greece
www.ics.forth.gr

SBA Research gGmbH
Floragasse 7, 1040 Wien, Austria
www.sba-research.org/

Fraunhofer ICT Group
Anna-Louisa-Karsch-Str. 2
10178 Berlin, Germany
www.iuk.fraunhofer.de

Eötvös Loránd Research Network
Számítástechnikai és Automatizálási Kutató Intézet
P.O. Box 63, H-1518 Budapest, Hungary
www.sztaki.hu/

INESC
c/o INESC Porto, Campus da FEUP,
Rua Dr. Roberto Frias, nº 378,
4200-465 Porto, Portugal
www.inesc.pt

University of Cyprus
P.O. Box 20537
1678 Nicosia, Cyprus
www.cs.ucy.ac.cy/

Institut National de Recherche en Informatique et en Automatique
B.P. 105, F-78153 Le Chesnay, France
www.inria.fr

Institute for Software Engineering and Software Technology "Jose María Troya Linero", University of Malaga
Calle Arquitecto Francisco Peñalosa, 18, 29010 Málaga
https://gp.uma.es/itis