Spring 2025 Newsletter

faith-ec-project.eu

Spring 2025

Follow Us

Follow us on social media for the latest updates, exclusive content, and more!

In this issue

In this edition, we explore key topics related to the **FAITH** project, including its latest developments, insights, and innovations. Below is the full index of this issue:

1. The vision of the project by Prof. Fotiadis, as the Coordinator 2
2. Description and vision of LSP5 - Wastewater management 3
3. Recent EC guidelines on the AI-ACT..... 5
4. Interview with Prof. Gregory Mentzas..... 6
5. Partner presentation: VEAS and CNR..... 8
6. News and events..... 10

Funded by the European Union under GA No 101135932. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Dimitrios Fotiadis

Project Coordinator

FIEEE, FEAMBES, FIAMBE

Prof. of Biomedical Engineering,
University of Ioannina

FORTH - Head of the
Unit of Medical Technology and
Intelligent Information Systems

Editor in Chief IEEE
Journal of Biomedical & Health Informatics

FAITH's vision

The **FAITH Project**[1] (Fostering Artificial Intelligence Trust for Humans towards the optimization of trustworthiness through large-scale pilots in critical domains, GA: 101135932) aims to address the awareness and cooperation of multi-disciplinary stakeholders who work on different aspects (e.g. technical, societal, legal, business, standards) of trustworthiness and different stages of the AI system's lifecycle. It aims to provide the practitioners and stakeholders of AI systems, not only with a comprehensive analysis of the foundations of AI trustworthiness, but also with an operational playbook for how to continuously assess, refine and build trustworthy AI systems. More specifically, the **FAITH AI_TAF** (trustworthiness assessment framework) develops a trustworthiness management methodology that identifies trustworthiness threats and vulnerabilities (not only technical but also social and human); evaluates risks and selects mitigation actions. The **FAITH AI_TAF** methodology follows a multiphase approach (6 phases), including: (i) cartography, (ii) threat analysis, (iii) impact assessment, (iv) vulnerability analysis, (v) risk analysis, and (vi) countermeasures.

Figure 1. The vision of FAITH

[1] https://faith-ec-project.eu/

To this end several other tools have been developed that support this methodology:
(i) the *FAITH AI TrustGuard* which is a checklist based risk assessment for AI-based systems in isolation,
(ii) the *FAITH AI TrustSense* which focuses on profiling the trustworthiness of the AI participant, and lastly
iii) the *FAITH AI Model Hub* which serves as a metadata collection repository for AI models and datasets integrating the notion of AI model passport and data passport.

The relevance, scalability, and impact of the FAITH AI_TAF will be assessed in **seven Large-Scale Pilots (LSPs)**. These pilots were chosen to reflect technological, societal, and regulatory challenges in AI deployment across the following critical domains:
(i) **LSP1 – Media** (AI-driven intelligent coaching application that will automatically detect disinformation and hate speech),
(ii) **LSP2 – Transportation** (AI-driven monitoring of the effectiveness of public transportation and the safety and security of passengers on board and in stations),
(iii) **LSP3 – Education** (Plato, an AI Learning Companion for STEM education, which supports students by providing automated guidance and feedback in STEM laboratory courses),
(iv) **LSP4 – Robotics/Drones** (AI driven maintenance of port infrastructure on data from underwater drones),
(v) **LSP5 – Industrial Processes** (Hybrid AI models for wastewater treatment),
(vi) **LSP6 – Healthcare** (AI-based automated prostate and zonal segmentation, visualization of the MR examination and the segmentations as an overlap, batch processing of the examinations), and
(vii) **LSP7 – Active Ageing** (AI-driven detection of behavioral patterns of elderly individuals from sensor data in their homes, alerting families about unusual bathroom duration times, indicating potential falls).

The FAITH AI_TAF will be evaluated through the proposed AI-based systems on the selected LSPs across different domains. Each LSP will produce domain-specific risk profiles of the FAITH AI_TAF.

AI in Wastewater Treatment

The FAITH Project's Large Scale Pilot

Asbjørn Følstad, SINTEF [asf@sintef.no]

In an ambitious stride towards sustainable innovation, the FAITH project advances a large scale pilot on the application of artificial intelligence (AI) in wastewater treatment. The pilot is hosted at Veas, Norway's largest wastewater treatment plant serving the greater Oslo area. Veas processes wastewater from an urban area with 800,000 residents, playing a vital role in maintaining the marine environment of the Oslo fjord. The pilot is coordinated by the research organization SINTEF that provides expertise in AI for industrial applications.

AI-powered innovations in Wastewater Treatment

The pilot focuses on optimizing the denitrification process, a crucial step in wastewater treatment where nitrate is converted to nitrogen gas through a biological process. Denitrification is important to preserve the ecological balance of the marine environment. By developing AI prediction models, the pilot aims to improve nitrogen removal and reduce usage of chemicals needed for denitrification, leading to substantial cost savings and environmental benefits.

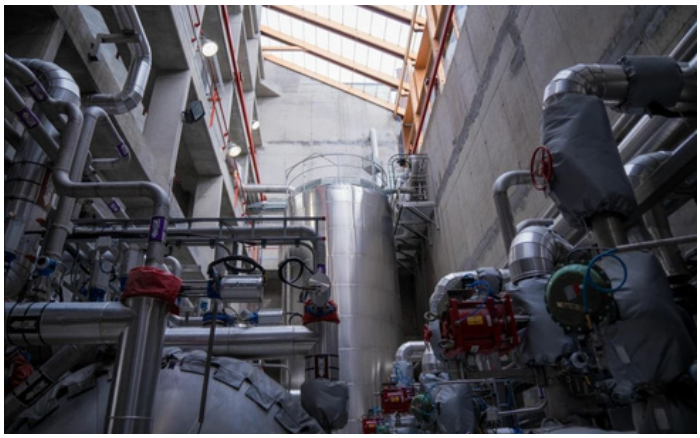

Figure 2. Vestfjorden Avløpsselskap (VEAS) operates the treatment plant at Slemmestad. It is Norway's largest, handling 100 million cubic meters of sewage per year. Photo by: Martin Leigland / TV2

The challenge of accurate and robust AI models

To leverage AI prediction models for denitrification of wastewater is a challenging task. Nitrate removal is a biological process whose state cannot be measured directly, and involves substantial seasonal and weather-driven variation. To address this, the SINTEF team explores different approaches to modelling, including initial forecasting models and chunk-based forecasting models where time series data are chunked in smaller time periods, and an ensemble of models trained on the individual chunks are used for model prediction. The best performing models will then be validated and deployed in the Veas process.

FAITH AI Trustworthiness Assessment Framework

At the heart of the FAITH project, is the AI Trustworthiness Assessment Framework (FAITH AI TAF), developed to validate and strengthen AI systems that require high levels of trustworthiness, such as those to be used in critical infrastructure as wastewater treatment. This framework will be applied to ensure the trustworthiness of the AI prediction models for Veas' wastewater treatment processes. The goal is to enhance the efficiency and effectiveness of the processes while ensuring that AI models are trustworthy and explainable.

Ensuring Trustworthiness and Stakeholder Involvement

Trustworthiness is a cornerstone of LSP5. The project emphasizes accuracy, robustness, interpretability, and human oversight of AI models. Regular stakeholder meetings, workshops, and direct interaction with AI predictions ensure that the models meet the needs and expectations of Veas personnel. The FAITH AI TAF will guide these assessments, ensuring the AI systems are reliable and transparent.

Looking Ahead

The FAITH large scale pilot of wastewater treatment has just completed its planning phase, and are now well into initial piloting. Here, the first AI models from the project are be made available to stakeholders at Veas, including process owners and engineers. In the third and final phase of the pilot, a replication pilot, stakeholder engagement will further expand and advanced AI models will be applied for process control provided successful initial trials and validation. The ultimate goal is to achieve significant improvements in nitrogen removal and cost efficiency setting a benchmark for wastewater treatment facilities worldwide.

By leveraging AI, the FAITH project and its partners, SINTEF and Veas, are paving the way for a cleaner, more sustainable future. This large scale pilot not only addresses critical environmental challenges but also showcases the potential of trustworthy AI in industrial applications.

Figure 3. Vestfjorden Avløpsselskap (VEAS)

Recent EC guidelines on the AI-ACT

Stefano Modafferi

IT Innovation Centre, University of Southampton [s.d.modafferi@soton.ac.uk]

The European Union's Artificial Intelligence Act (AI Act) establishes the first comprehensive legal framework for regulating AI systems, adopting a risk-based approach to ensure safety, fundamental rights, and trustworthy AI development. To support the effective implementation of its provisions, the European Commission (EC) issued two key guidelines in February 2025: one on the definition of AI systems (February 6) and another on prohibited AI practices (February 4). These non-binding guidelines aim to clarify critical aspects of the Act, assisting stakeholders in navigating compliance requirements.

Guidelines on the Definition of AI Systems (Feb.6, 2025)

This guideline clarifies what constitutes an "AI system" under Article 3 of the EU AI Act. Its purpose is to ensure that stakeholders can accurately identify whether their software or technology falls within the scope of the Act. Key takeaways include:

- **Broad and Technology-Neutral Definition:** An AI system is defined as a machine-based system operating with varying levels of autonomy to generate outputs such as predictions, recommendations, or decisions that influence physical or virtual environments.
- **Case-by-Case Analysis:** Determining whether a system qualifies as an AI system requires analysis based on functionality rather than specific technologies. This ensures flexibility and adaptability to emerging technologies.
- Evolving Framework: The definition is designed to accommodate future advancements in AI, ensuring that the regulatory framework remains relevant over time.

These guidelines provide foundational clarity for businesses and developers to understand whether their systems are subject to regulation under the AI Act. For more details, refer to the official EC document: **Guidelines on the definition of an artificial intelligence system** [2]

Guidelines on Prohibited AI Practices (Feb. 4, 2025)

This set of non-mandatory guidelines elaborates on Article 5 of the EU AI Act, which lists practices deemed to pose "unacceptable risks" and are therefore prohibited. Key takeaways are:

- The Draft Guidelines aim to increase clarity and provide insight into the Commission's interpretation of Prohibited Practices under Article 5 of the Act.
- The Draft Guidelines are lengthy but are still in draft form and, even when finalized, will be non-binding. All guidance provided therein is subject to the formal requirements set forth in the Act.
- Though the Draft Guidelines are not comprehensive, they are a helpful step in assessing whether an AI System qualifies as prohibited under the Act.

These include applications that violate fundamental rights or EU values. Key examples include:

- Subliminal or Manipulative Techniques: Prohibits AI systems that use subliminal techniques or manipulative strategies beyond individuals' awareness to distort their behavior in ways that cause significant harm.
- Exploitation of Vulnerabilities: Bans systems exploiting vulnerabilities related to age, disability, or socio-economic conditions to distort behavior and cause harm.
- Biometric Categorization: Outlaws systems inferring sensitive attributes (e.g., political beliefs or sexual orientation) from biometric data unless strictly necessary for law enforcement with safeguards.

[2] https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act

- **Real-Time Remote Biometric Identification:** Prohibits its use in publicly accessible spaces except for narrowly defined law enforcement purposes with judicial authorization.
- **Untargeted Facial Scraping:** Prohibits indiscriminate scraping of facial images from online platforms or CCTV footage to create facial recognition databases without individuals' consent. Examples include commercial facial recognition systems built using images scraped from social media platforms.
- **Emotion Recognition in Sensitive Contexts:** Bans AI systems designed to infer emotions in workplaces and educational institutions due to concerns about their reliability and potential for discriminatory outcomes.
- **AI for Social Scoring:** Prohibits systems evaluating individuals' behaviour or characteristics unrelated to context, leading to unjustified discrimination or exclusion.

For further information, consult the official EC publication: Guidelines on prohibited artificial intelligence practices.

Conclusion

Together, these two guidelines provide essential clarity for implementing the EU AI Act's provisions. By defining what constitutes an AI system and detailing prohibited practices, they help ensure consistency across Member States while fostering ethical innovation and safeguarding fundamental rights. These resources are invaluable for businesses seeking to align their operations with the EU's regulatory framework for AI.

An interview with Prof. Gregory Mentzas

[gmentzas@mail.ntua.gr]

Why is FAITH considered a straightforward approach to trustworthy AI?

AI is a key driver for digital transformation, but regulatory and societal scrutiny is required to address and mitigate related risks and guarantee trustworthiness of AI models and applications. While traditional AI evaluation focuses on accuracy, trustworthiness also depends on transparency, security, fairness, and privacy. Ensuring AI trustworthiness requires a lifecycle approach, from data preparation to deployment and monitoring. Key risks include issues like bias, outdated datasets, model unpredictability and others, requiring a balance between privacy, explainability, and fairness.

FAITH will introduce a human-centric Trustworthiness Assessment Framework (FAITH AI_TAF), that integrates regulatory requirements, risk management, and real-world validation. Large-scale pilots across seven critical sectors (e.g., healthcare, media, robotics) will refine AI trustworthiness measures and promote cross-domain standardization. By extending risk management to socio-technical AI challenges, FAITH will enhance AI governance, supporting the development of ethical, legally compliant, and trustworthy AI systems.

What makes FAITH different from other AI trustworthiness initiatives in Europe?

Prof. Gregory Mentzas

Gregoris Mentzas is full Professor of Management Information Systems, School of Electrical and Computer Engineering, National Technical University of Athens (NTUA) and Founder and Director of the Information Management Unit (IMU) at the Institute of Communication and Computer Systems (ICCS)

FAITH brings two key innovations beyond the current state of the art. First, each large-scale pilot will enhance existing methods at both technical and organizational levels within the human-centric, risk-management-driven FAITH AI_TAF framework. Technical advancements will improve the performance of ML methods, such as real-time fleet assessment in the transport pilot. Organizational improvements will involve defining trustworthy AI requirements, categorizing threats, implementing risk mitigation measures, and leveraging FAITH tools like the AI Model Passport. This will enable the delivery of robust, trustworthy AI systems.

Second, the FAITH AI_TAF itself will evolve through the integration of innovative tools, including the AI Model Passport and System Trust Modeller (STM). Pilot projects will refine the FAITH framework by providing domain-specific insights into trustworthiness needs and threats. By incorporating these insights into a structured risk analysis approach, FAITH AI_TAF will provide an actionable framework of tools and guidelines, supporting the development of transparent, traceable, and trustworthy AI across critical sectors.

How will FAITH's solution be useful for other research and innovation projects?

The risk-based FAITH AI_TAF could serve as a foundational tool for future research projects across various domains. By establishing a structured approach to assessing risks related to AI reliability, fairness, transparency, and security, this framework can be adapted for different industries. Future research can build upon this framework to develop domain-specific trust assessment methodologies, integrating industry standards and regulatory requirements to ensure compliance and ethical deployment of AI technologies. Additionally, the framework can be expanded to facilitate interdisciplinary research on AI ethics and governance. Researchers studying human-AI interactions, algorithmic bias, or explainability can leverage this structured process to quantify and compare trustworthiness metrics across different AI models. This can lead to more standardized benchmarks for evaluating AI safety and effectiveness, supporting collaborative efforts between academia, industry, and policymakers. By incorporating evolving risk factors, such as adversarial attacks or data shifts, future studies can refine and enhance AI trust validation techniques to address emerging threats. Beyond technical validation, the framework can be instrumental in researching how public trust in AI can be fostered. Researchers in social sciences and behavioral studies can use it to analyze how different risk mitigation strategies impact user perceptions of AI systems. This can contribute to developing guidelines for responsible AI communication and user-centered design. Moreover, as AI regulations continue to evolve globally, the framework can support research in legal and compliance domain
ensuring that AI trustworthiness assessments align with emerging policies and standards. This multidisciplinary applicability ensures that the framework could remain a valuable asset in the pursuit of responsible and ethical AI development.

Example, how will the LSP in education be benefited by the FAITH AI_TAF?

Current science education reforms emphasize key competencies and inquiry-based learning, but high-stakes exams still dominate, reinforcing traditional teaching and assessment methods. These methods fail to evaluate students' ability to think scientifically, solve complex problems, and engage in authentic STEM learning. AI-driven assessment tools offer a potential solution, continuously monitoring student progress, providing feedback, and supporting personalized learning paths.

The education Large Scale Pilot (LSP) of FAITH will evaluate the trustworthiness of AI-based student assessment in inquiry-based laboratory work. Our team at the Information Management unit of ICCS together with the Ellinogermaniki Agogi school, have developed an AI Learning Companion that provides real-time guidance, corrects misconceptions, and adapts to students' needs, allowing teachers to act as facilitators rather than traditional instructors.

Within the education LSP we will be using and adapting the FAITH AI_TAF to ensure the trustworthiness of the AI Learning Companion and address concerns about efficacy, explainability, ethical implications, and the impact on teachers. This pilot will address these challenges through large-scale implementation and adoption of the AI_TAF. Our pilot will involve 40 teachers and 1,000 students (ages 12–15) across multiple schools. Ultimately, this use case aims to validate the trustworthiness of AI-driven assessment as a sustainable, effective alternative to traditional student evaluation methods in STEM education.

Partner Presentation

ISTI-CNR

Italy
www.cnr.it

Consiglio Nazionale delle Ricerche

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo"

ISTI is the largest institute within the National Research Council (CNR) dedicated to Computer Science. It boasts a team of approximately 170 permanent staff members and 100 fellows or PhD students. Committed to scientific excellence, ISTI plays a key role in advancing research and fostering technology transfer. The accessibility. With a team of approximately thirty researchers and technologists, the SI Lab brings together expertise in computer science, engineering, physics, and mathematics. It operates dynamically within national and international academic and industrial networks, fostering innovation in automated vision and information technologies. The lab's research spans multiple application areas, including health and well-being, medical imaging, sustainable agriculture, cultural heritage, biodiversity and environmental protection, and sustainable mobility. By leveraging AI-driven solutions and advanced signal processing techniques, SI Lab aims to develop innovative methodologies to

ISTI unit involved in this project is represented by the "Signals & Images" (SI) Lab. The lab specializes in computer vision, artificial intelligence, intelligent systems, statistical signal processing, bio-inspired mathematical modeling, uncertainty quantification, topological data analysis, human-computer interaction, and also address pressing environmental challenges and promote biodiversity conservation.

In the FAITH project, ISTI is contributing to the design and development of the FAITH AI_TAF and to the definition of an AI Model Passport. In addition to its role in the Large-Scale Pilot on Healthcare (LSP6), ISTI serves as the work package leader for Large-Scale Pilot on public transportation (LSP2), carried out in collaboration with MERMEC Engineering. LSP2 activities include designing, developing, and assessing the trustworthiness of computer vision systems for monitoring carriages and other areas of interest in the railway sector.

Partner Presentation

Veas

Norway
veas.nu

Leading the Way in Wastewater Treatment Innovation

Veas, the largest wastewater treatment plant in Norway, is a cornerstone of the FAITH project, bringing its extensive expertise and innovative spirit to the forefront of this groundbreaking initiative. Owned by the municipalities of Oslo, Bærum, and Asker, Veas processes approximately 100 million cubic meters of wastewater annually, serving around 800,000 residents in the greater Oslo area.

Located by the inner Oslo fjord, Veas is not just a wastewater treatment facility; it is a beacon of innovation and sustainability. The plant plays a crucial role in maintaining the marine environment of the Oslo fjord, which is under significant ecological pressure. Veas is committed to exceeding regulatory requirements by aiming for 80% nitrogen removal, well above the mandated 70%. Veas' ambition to be a lighthouse in the sector is evident in its strategic focus on digital technology and AI. The plant's fully automated processes generate vast amounts of data, which are harnessed to optimize operations and reduce costs. Through the FAITH project, Veas aim to make use of enhanced AI-based decision support systems to strengthen the efficiency and effectiveness of wastewater treatment. For this, the trustworthiness of the AI system is key, so as to reap the benefits of AI while allowing process owners and operators to be in full control.

By leveraging cutting-edge technology and fostering a culture of continuous improvement, Veas is setting new standards in wastewater management. Its participation in the FAITH project underscores its commitment to innovation and its role as a leader in the application of trustworthy AI in industrial processes.

News and events

The first stakeholder workshop in Pisa

On February 21st, 2025, the first FAITH stakeholder workshop took place at the CNR premises in Pisa, bringing together a dynamic mix of experts and industry leaders to discuss the future of AI-driven public transportation solutions.

With around 20 in-person participants, the event featured representatives from both private and public transportation sectors, alongside legal and scientific experts. This carefully curated group ensured a diverse and balanced exchange of ideas, fostering meaningful discussions on trustworthiness and innovation in AI applications.

The success of the event was greatly reinforced by the participation of the coordinator of DITECFER (District for Railway Technologies, High Speed, Networks' Safety & Security), who not only joined the meeting but also facilitated the involvement of key stakeholders in the railway industry.

This first event did not include other categories of stakeholders—such as end users, onboard train staff, and railway service personnel—who are highly relevant to the objectives of the Large Scale Pilot on public transport. However, specific participation strategies will be implemented to engage them in future initiatives.

The first stakeholder workshop primarily targeted technology and innovation actors, as well as policymakers. More specifically, it was attended by public entities, national transport operators, certification bodies, Industry companies and legal and scientific experts.

During the event, various issues were discused like the impact of the **AI Act** on innovation in the railway sector, the objectives of **FAITH** for the transportation domain and a focus group was designed on the concept of trustworthiness in the railway sector and its main challenges.

FAITH At Madeira Digital Transformation Week 2024

Prof. Dimitris Fotiadis, represented the project at Madeira Digital Transformation Week 2024, held from June 20–24. During the event, Prof. Fotiadis delivered a presentation titled "Enhancing Trust in AI Across Critical Application Domains," highlighting the pressing need for trustworthy AI in high-stakes sectors. For this reason, FAITH project is being validated through large-scale pilots across key sectors, demonstrating how trustworthy AI can be practically achieved and sustained in real-world environments.

FAITH at EMRA 2024 in Arenzano

The FAITH project was presented at the Workshop on EU-funded Marine Robotics and Applications (EMRA) 2024, held in Arenzano, Italy, from May 27–29. The event was hosted by the Institute of Marine Engineering of the Italian National Research Council (INM-CNR), EMRA is a key event in the marine robotics and applications community.

Representing FAITH, Mark Tanner (Project Manager of Innovation & Port Cluster Development) and Rafael Company (Director of Safety and Security), both from Fundación Valenciaport (VPF), shared insights into the project's objectives and progress. Their presentation offered an overview of FAITH and a focused look at its Large-Scale Pilot (LSP) activities.

EMRA 2024 brought together experts from H2020 and Horizon Europe projects, along with representatives from industry, end-user communities, and research institutions. The workshop provided a vibrant platform for interdisciplinary exchange and collaboration on marine robotics and enabling technologies—making it an ideal setting to introduce FAITH's contributions to AI trustworthiness in maritime applications.

ATC Presents FAITH Project at EBDVF 2024

ATC's A. Ramfos took part in a high-level panel during the European Big Data Value Forum (EBDVF) 2024, held in Budapest from October 2–4. Speaking in the session titled "RISK assessment and trust in socio-technical AI systems," Ramfos represented the FAITH project and shared insights on how a structured risk management approach can help ensure trustworthiness in AI systems. As highlighted by the event organizers, with AI rapidly integrating into various sectors, addressing the challenges of risk and trust is more critical than ever. The session explored the complexities of socio-technical AI systems, where human, organizational, and technological dimensions intersect—bringing both challenges and opportunities for building reliable, ethical AI.

More news

Publications

- Kitty Kioskli ,Eleni Seralidou and Nineta Polemi, **A Practical Human-Centric Risk Management (HRM)** Methodology,Electronics 2025, 14(3), 486; https://doi.org/10.3390/electronics14030486
- Nikolaos Antonios Grammatikos, Evangelia Anagnostopoulou, Dimitris Apostolou and Gregoris Mentzas **An AI-Powered Learning Companion for Adaptive and Personalized STEM Education,** Conversations 2024 workshop
- Kitty Kioskli, Antonios Ramfos, Steve Taylor, Leandros Maglaras, Ricardo Lugo, **Optimizing AI System Security: An Ecosystem Recommendation to Socio-Technical Risk Management**, AHFE 2024 Hawaii Conference https://doi.org/10.54941/ahfe1005635
- Pål V. Johnsen, Eivind Bøhn, Sølve Eidnes , Filippo Remonato, Signe Riemer-Sørensen, **Recency-Weighted Temporally-Segmented Ensemble for Time-Series Modeling** arXiv:2403.02150 https://doi.org/10.48550/arXiv.2403.02150

More publications

@faith_ec_project