European Partnership on Metrology
Call 2025 – Health, Integrated European Metrology, Regulation and Research Potential

Selected Research Topic number: **SRT-s20**
Version: 1.0

Title: Metrological foundations for benchmarking the performance of Artificial Intelligence

Abstract

Proposals addressing this topic should develop harmonised methodologies to measure and quantify the uncertainty associated with outputs of AI algorithms. Based on case studies from applications in AI medicine, guidelines will be established for transferring robust uncertainty quantification approaches to high-risk AI applications and fields of metrology. This will enhance the robustness and transparency of AI-based decision-making processes and support future extensions of the established uncertainty assessment framework.

Keywords

Machine learning, uncertainty assessment, trustworthy artificial intelligence, AI in metrology, AI in medicine, standardised datasets.

Background to the Metrological Challenges

Artificial Intelligence (AI) has made tremendous progress in recent years and has become one of the key drivers for innovations in economy, medicine, and daily life. AI promises more cost-effectiveness in industrial production, novel approaches to production of goods, such as advanced manufacturing, as well as better quality of products, and their maintenance, during their life cycle. In medicine the introduction of AI in clinical practise, not only makes workflows more efficient, but also appears to be a game changer for the development of automated diagnosis and personalised medicine.

European Union (EU) has initiated regulation of AI by the Artificial Intelligence Act (AI Act), to support innovation while protecting its citizens. Building trust into AI appears to be a prerequisite for exploiting the potential of AI for Europe's economic competitiveness. The AI Act has been complemented by a standardisation request to CEN/CENELEC, aimed at the development of harmonised standards for quality assurance and conformity assessment of AI and AI-based product and devices. While the development of standards addressing governance, quality and risk management related to AI as a software and to testing of AI tools are in progress, the development of standards on technical aspects of AI testing is hampered by the lack of generally accepted metrics, for measuring AI performance and quality of data used for training and testing.

Successfully addressing the proposed topic, will enable an extension of the established framework for uncertainty specification in metrology, according to the GUM, to include also additional uncertainties coming from the AI models (epistemic) and the data (aleatoric) used in their training and validation. This will lay the foundations for quantitatively measuring AI performance in certification, in a unique and consistent manner. Beyond fulfilling the requirements from the EU AI Act to develop a metrology basis for the measurement of AI performance, the project will thus help build trust into AI-empowered measurement devices, as well as into applications of AI in many fields identified as high-risk areas.

Objectives

Proposers should address the objectives stated below, which are based on the PRT submissions. Proposers may identify amendments to the objectives or choose to address a subset of them in order to maximise the

EURAMET MSU,
Email: part.msu@euramet.org
Metpart.eu

European Partnership Co-funded by the European Union
METROLOGY PARTNERSHIP EURAMET

overall impact, or address budgetary or scientific / technical constraints, but the reasons for this should be clearly stated in the protocol.

The proposal shall focus on the traceable measurement and characterisation of uncertainty associated with the outputs of AI algorithms.

The specific objectives are

1. To develop a harmonised and comprehensive framework for assessing the uncertainty of the output from AI algorithms and the quality of data used in their training and validation, demonstrating its applicability with at least three use cases from medical or other high-risk applications of AI (e.g. dose calculation in radiotherapy, autonomous driving) and applications of AI in metrology or metrological services (e.g. for environmental or infrastructure monitoring).
2. Using outputs of Objective 1, to develop guidelines for the assessment of these AI-specific uncertainties in the above mentioned and other application areas in metrology and to develop open-source software tools to support the dissemination and implementation of the proposed framework, paving the way to a future extension of the established uncertainty assessment framework provided by the JCGM Guide to the Expression of Uncertainty in Measurements.
3. To develop quantitative metrics for assessment of explainability, accuracy and robustness in AI outputs and their derived decision; to develop metrological validation methods for AI/ML systems by establishing benchmark procedures using standardised real-world datasets for targeted applications in key areas such as healthcare, infrastructure monitoring, environmental management and industrial control.
4. To formulate a generic methodology (e.g., item-response logistic regression) for assessing decision risks in conformity assessment and AI classification tasks, and to propose strategies to take the results of uncertainty assessments into account for improved decision-making, particularly in critical fields like healthcare diagnostics or precision measurements.
5. To demonstrate the establishment of an integrated European metrology infrastructure and to facilitate the take up of the technology and measurement infrastructure developed in the project by the standards developing organisations (ISO/IEC 22989, ISO/TC215, ISO/IEC JTC1 / SC42, CEN/CENELEC JTC 21) and end users (healthcare institutions, environmental and industrial organisations).

These objectives will require large-scale approaches that are beyond the capabilities of single National Metrology Institutes and Designated Institutes. To enhance the impact of the research work, the involvement of the larger community of metrology R&D resources both within and outside Europe, plus engagement with existing European research infrastructures and European Partnerships is recommended. A strong industry involvement is expected in order to align the project with their needs and guarantee an efficient knowledge transfer into industry and end users. Where relevant, proposals are encouraged to build on, or seek collaboration with, existing projects and develop synergies with other relevant European, national or regional initiatives and funding programmes. In particular, links are encouraged with (a) the projects funded under earlier relevant topics of the Horizon Europe programme; or (ii) other relevant European Partnerships such as ADRA.

Proposers should establish the current state of the art and explain how their proposed project goes beyond this. In particular, proposers should outline the achievements of the EMPIR project 18HLT07 Medalcare, and the Metrology Partnership projects 22HLT01 QUMPHY, 23NRM03 BioAirMet, and 22HLT05 MAIBAI.

Proposers should note that the programme funds the activity of researchers to develop the capability, not the required infrastructure and capital equipment, which must be provided from other sources.

EURAMET expects the average EU Contribution for the selected JRPs in this TP to be 2.1 M€ and has defined an upper limit of 2.6 M€ for this proposal.

EURAMET also expects the EU Contribution to the external funded beneficiaries to not exceed 25 % of the total EU Contribution across all selected projects in this TP.

Any industrial beneficiaries that will receive significant benefit from the results of the proposed project are expected to be beneficiaries without receiving funding or associated partners.

Potential Impact

Proposals must demonstrate adequate and appropriate participation/links to the ’end user’ community, describing how the project partners will engage with relevant communities during the project to facilitate

knowledge transfer and accelerate the uptake of project outputs. Evidence of support from the “end user” community (e.g. letters of support) is also encouraged.

You should detail how your proposal’s results are going to:

- Address the SRT objectives and deliver solutions to the documented needs,
- Feed into the development of urgent documentary standards through appropriate standards bodies,
- Facilitate improved industrial capability, or improved quality of life for European citizens in terms of personal health, protection of the environment and the climate, or energy security,
- Transfer knowledge to the healthcare and environmental sectors.

You should detail other impacts of your proposed JRP as specified in the document “Guide 4: Writing Joint Research Projects (JRPs)”

You should also detail how your approach to realising the objectives will further the aim of the Metrology Partnership to develop a coherent approach at the European level in the field of metrology and include the best available contributions from across the metrology community. Specifically, the opportunities for:

- improvement of the efficiency of use of available resources to better meet metrological needs and to assure the traceability of national standards
- the metrology capacity of EURAMET Member States whose metrology programmes are at an early stage of development to be increased
- organisations other than NMIs and DIs to be involved in the work.

Timescale

The project should be of up to 3 years duration.

Additional information

The links provided in this section are only correct at the time of publication up until the end of the Call year.

These references have been provided by EURAMET.

[1] EMN for Mathematics and Statistics Strategic Research Agenda
https://www.euramet.org/european-metrology-networks/mathmet/strategy/strategic-research-agenda

[2] EMN for Advanced Manufacturing
https://www.euramet.org/european-metrology-networks/advanced-manufacturing/strategy/strategic-research-agenda