



CHIST-ERA CALL 2020 – OPEN SCIENCE

OPEN DATA FOR EVALUATION AND CERTIFICATION

DR. AGNES DELABORDE
agnes.delaborde@lne.fr

01/10/2020



ABOUT LNE

FRENCH NATIONAL LABORATORY FOR METROLOGY AND TESTING



A RECOGNIZED AND MULTI-DOMAIN ORGANISM



Transportation



Medical / health



Energy



Institutional
authorities



Defence



Industry



Construction

Notified for more than 20
European Directives

Autorised by national
accreditation organism (COFRAC)

Calibration

Testing

Certification
(products, management systems)

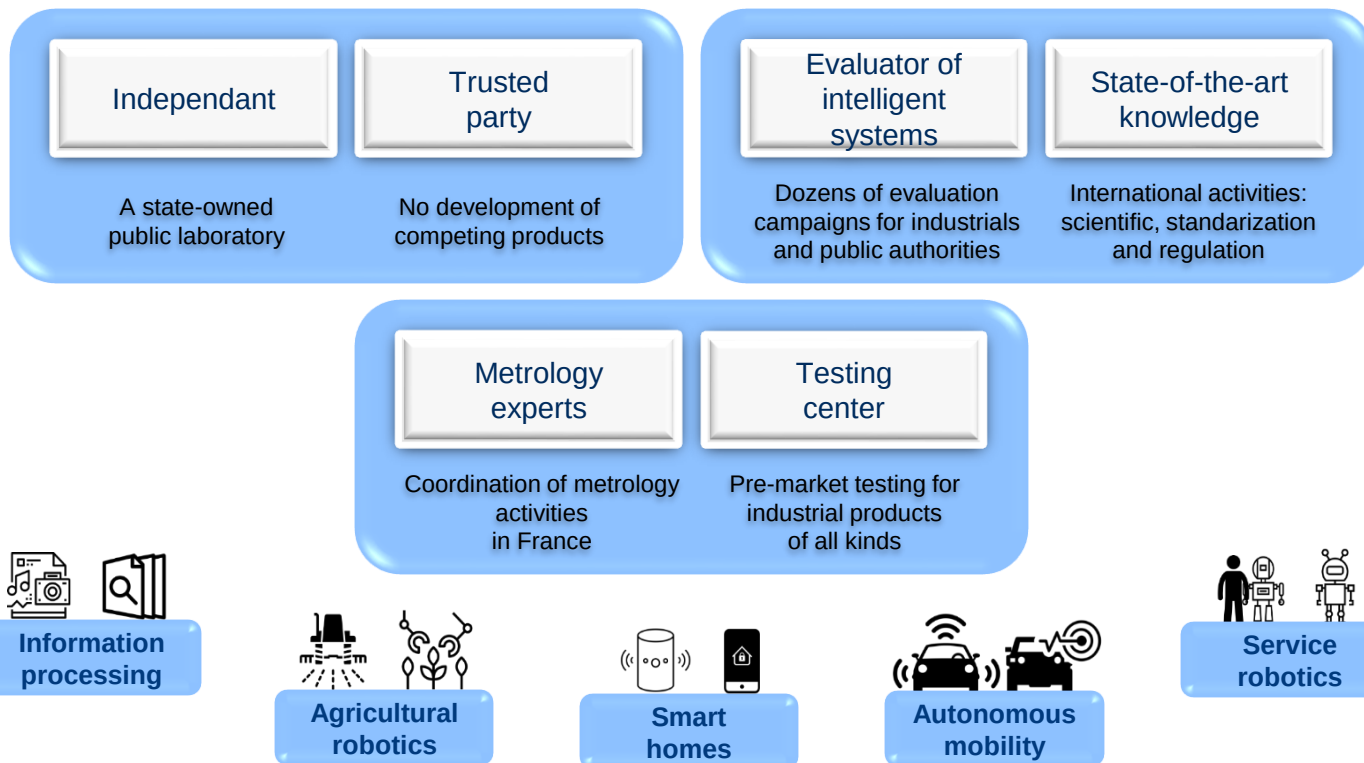


Instrumentation



Consumables

EXPERTISE: EVALUATION OF ARTIFICIAL INTELLIGENCE



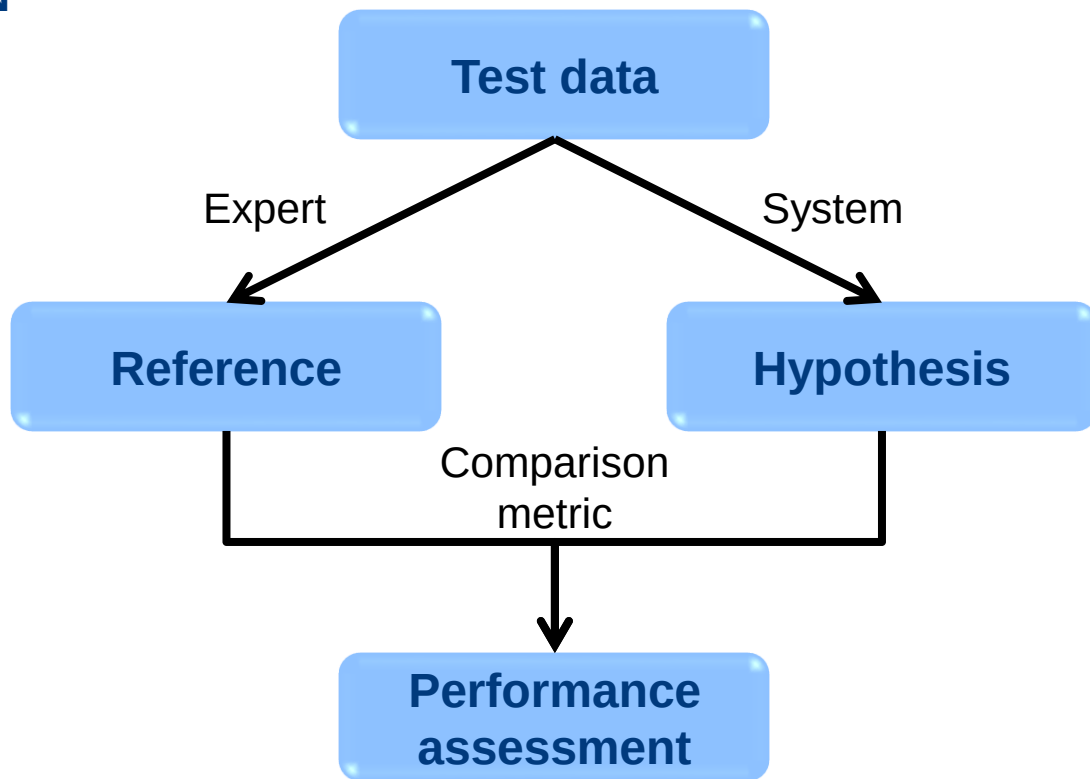


EVALUATION AND CERTIFICATION

WHY IS EVALUATION SO IMPORTANT?

- Performance
 - More than a competitiveness argument: a regulatory requirement (safety, explainability, traceability)
- Towards certification
 - To ensure the reliability of deployed AI systems in Europe
 - Based on an evaluation on representative data

EVALUATION



CERTIFICATION

An ongoing work

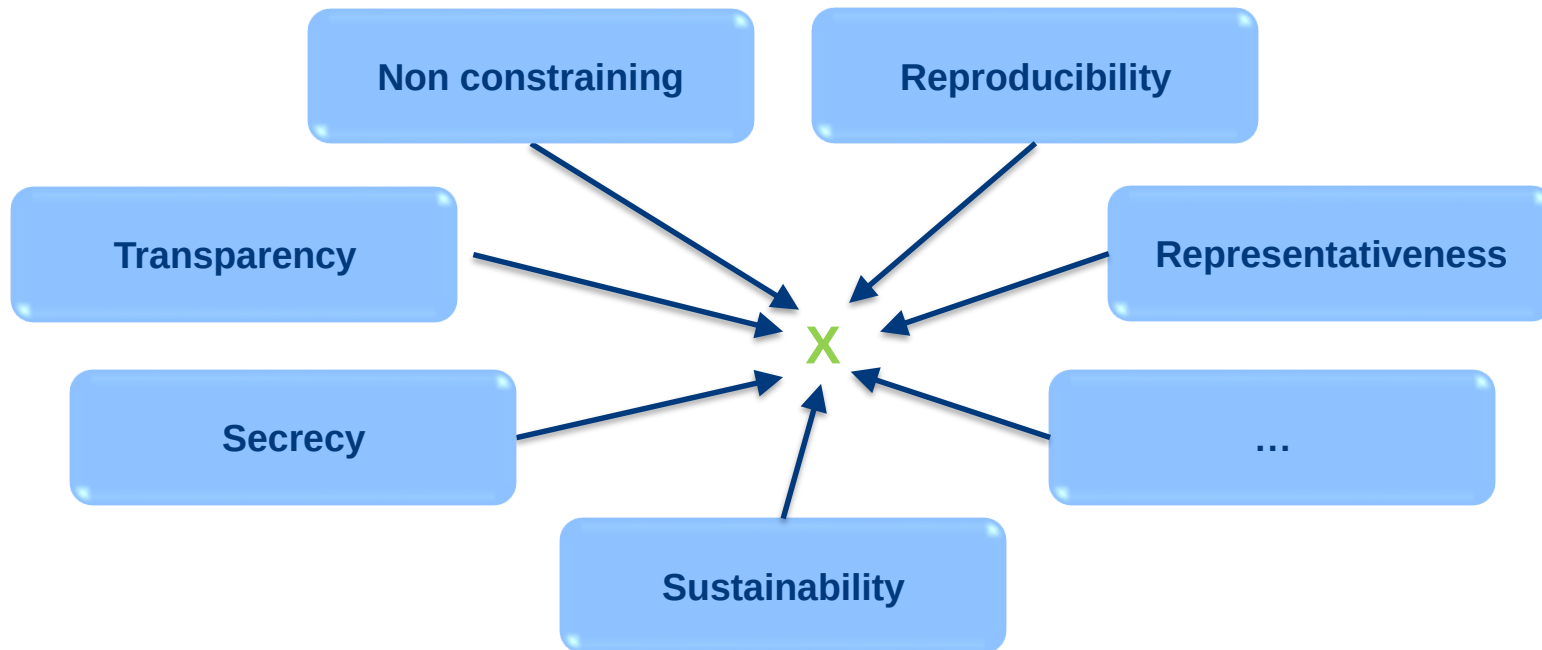
- What certification?
 - **Product**: the behaviour of the AI software module is compliant (tests on the module) and/or
 - **Process**: the module has been correctly designed (checking of learning/development phases)
- What references?
 - **Reference values** (minimal error rates, ground truth, etc.)
 - Testing **methods**, metrics
 - Testing **environments** (simulation, databases, etc.)
- Is consensus reachable?
 - **Completeness and sustainability** of testing databases (for which algorithm, which application?)
 - Ensure the **non-distribution** of the test database



THE OBJECTIVES

TOWARDS A TRADE-OFF

Achieve convergence between certification requirements and data regulations



OPEN DATA FOR EVALUATION AND CERTIFICATION

- Designing open access models of test data, that will allow the collection of adapted data sets for the evaluation in the context of certification, re-certification (evolving systems) and self-certification
- A model of governance of testing data to enhance the interactions between industry, public authorities and research

Thank you for your attention

agnes.delaborde@lne.fr

ABSTRACT OF THE TALK

Open data for evaluation and certification

The certification of data-processing systems (including artificial intelligence systems) is of paramount importance to guarantee their regulatory compliance in Europe. The definition of certification strategies is a current topic, and one can expect that a part of the certification procedures will rely on the evaluation on representative data. However, the data on which compliance tests have to be performed is still an issue for public authorities, and research has not yet produced a sufficient answer. These data need to be conceived in an “open data” spirit, since the process must be reasonably transparent, and industry must be aware of the nature of the compliance tests (in the case of certification by an accredited third party) or even be able to perform these tests (for self-certification in non-critical domains). Evaluators must then have test datasets at their disposal. However, there are obvious limitations to creating evaluation data sets for certification: how can one collect datasets that are sufficiently representative of the variety of systems to be evaluated, unknown to the developers (to prevent training), that do not hinder innovation by imposing too many constraints on the developers, based on data both collected from sources sufficiently different from those used for learning, but also sufficiently similar to ensure comparability... Building and storing pools of test datasets for certification seems to be a really limited option. We envision one possible solution: designing open access models of test data, that will allow the collection of adapted data sets for the evaluation in the context of certification, re-certification (evolving systems) and self-certification. Europe needs to be structured to be able to face the fast evolution of data technologies; a model of governance should also be designed to enhance the interactions between industry, public authorities and research.