Combining Data-Driven and Knowledge-Based AI Paradigms for Engineering AI-Based Safety-Critical Systems

Engineering AI-based critical system induces various challenges

Data & Knowledge Engineering
- Feature characterization
- Data & Knowledge quality
- Representativeness
- Corpus balancing & biases reduction

Algorithm Engineering
- Specifiability
- Traceability
- Correctness / Validity
- Accuracy
- Complexity
- Transparency
- Vulnerability

Human-AI Interaction
- Usability
- Interpretability / Explainability
- Human-AI dialogue
- Ethics by design

Safety & Cybersecurity
- Provability
- Verifiability (test)
- Robustness
- Integrity / Resilience

Software & System Engineering
- Repeatability
- Performance
- Maintainability
- Auditability
- Monitorability

A 1st end-to-end framework to assess Trust through Risk Mgmt & Assurance Case

Trustworthy attributes definition & associated KPIs

Revisiting all engineering disciplines to propose a sound deployment of AI components within safety-critical systems

Juliette MATTIOLI (Thales)
Gabriel PEDROZA (CEA)
Souhaiel KHALFAOUI (Valeo & IRT SystemX)
Bertrand LEROY (Renault & IRT SystemX)