

RESPONSIBLE AI

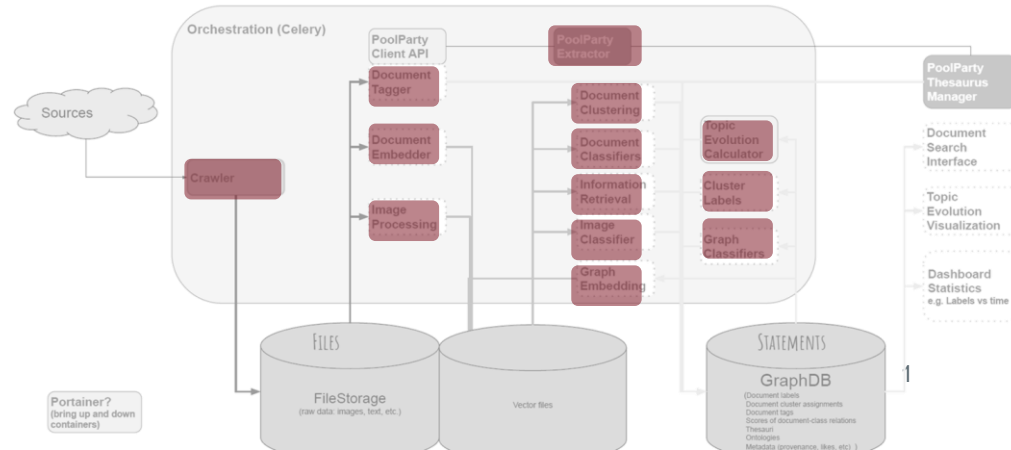
Role and Responsibility of an applied research institution

- Projects with public authorities
- Decision support systems

- Systems are getting more complex
 - **Past:** One-project – one-model / One-problem – one-model
 - **Presence / Future:** One-problem – many-models / One-project – a lot of models

- Challenge:
 - Integrated consortial projects

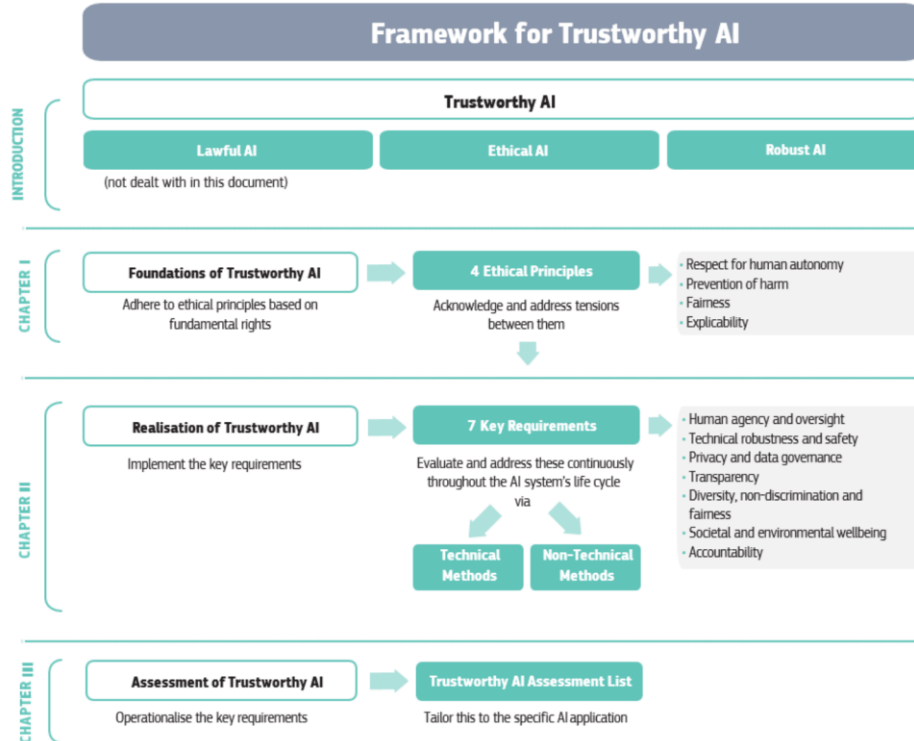
AI Models – Project RAIDAR (Detect Hate Speech & Radicalization)



TRUSTWORTHY AI

EU Ethics Guidelines for trustworthy AI

Framework for Trustworthy AI



7 Requirements of Trustworthy AI



TRUSTWORTHY AI @ AIT

Ethical AI

- AIT AI-Ethics Lab (P. Biegelbauer)

Robust AI

- Data Literacy Principles
- Software / Model Quality (ML-Ops)
- Implementing Standards

Legal AI

- Data Protection Officer (GDPR compliance)
- Research Projects
 - RAIDAR: Lawfulness by Design (by Research Institute)

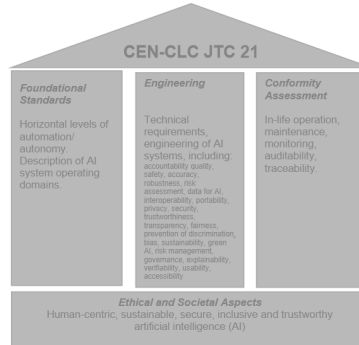
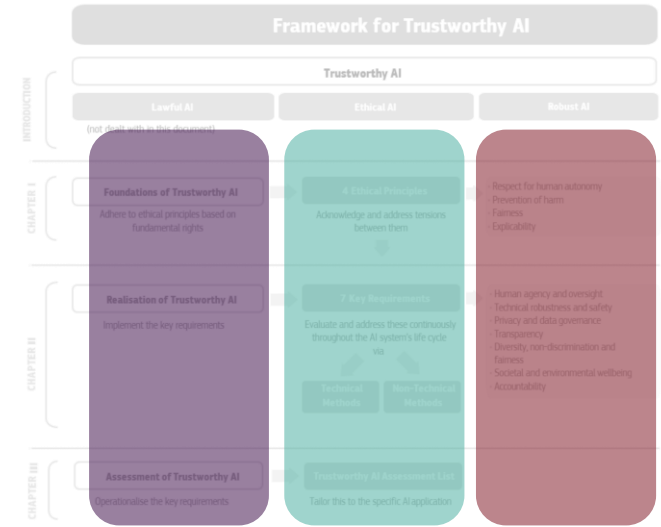


Figure 3: Overview of JTC 21 framework



Ethics guidelines for trustworthy AI

STANDARDIZATION

and implementing standards

- **Participation: Austrian Standards – WG Artificial Intelligence**

- accountability quality
 - Safety
 - accuracy
 - robustness
 - Risk assessment / management
 - Interoperability
 - portability
 - Privacy
 - security
 - Transparency
 - fairness
 - prevention of discrimination
- Bias
 - Sustainability / green AI
 - Governance
 - explainability
 - Verifiability
 - usability
 - Accessibility

Implemented (company/center level)

Projects

e.g., STARLIGHT, Green AI projects, DeepRUL, RAIDAR, defalsif-ai

Thematic Focus / Strategy 22/23

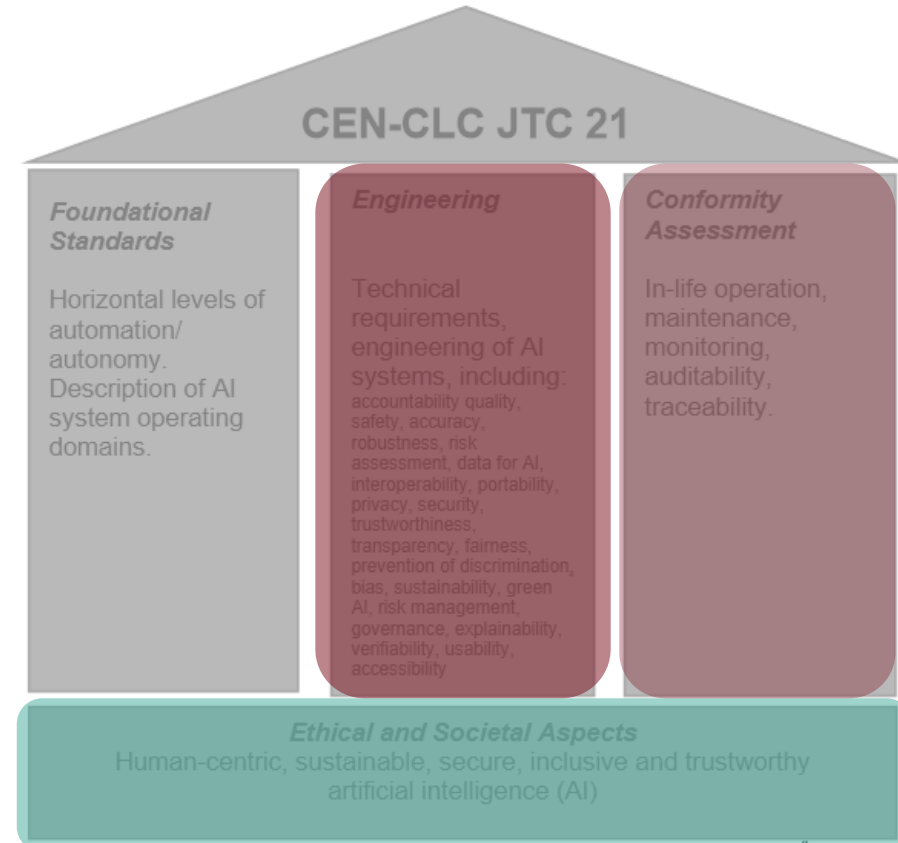


Figure 3: Overview of JTC 21 framework