Bundesnetzagentur

European standardisation synergy: towards the AI Act

Taras Holoyad
Federal Network Agency

ETSI MTS #90, Mainz, 26.-27.09.2023

European AI standardisation

- AI Act: state of play
- Standardisation Request (SR) on AI
- Recent Work Items relevant for SR
- European cooperation: ETSI & CEN-CENELEC

AI Act: state of play (1/2)

Current EU AI Act trilogue: main areas

- Definition of AI
- Prohibited AI systems
- Requirements for High-risk AI systems
- Requirements for foundation models
- Enforcement for non-compliance with legislation

Source: iapp - Contentious areas in the EU AI Act trilogues, https://iapp.org/news/a/contentious-areas-in-the-eu-ai-act-trilogues

AI Act: state of play (2/2)

European AI Act: temporal development

Now	Trilogue: Commission, Parliament & Council
Late **2023** (expected)	Political agreement on the AI Act is reached.
Early **2024** (expected)	The finalized AI Act is adopted.
Late **2025** - Early **2026** (expected)	Following a likely 18-24-month transition period, the AI Act comes into effect.

Source: iapp - Contentious areas in the EU AI Act trilogues, https://iapp.org/news/a/contentious-areas-in-the-eu-ai-act-trilogues

Standardisation Request on AI (draft): 10 Topics for HEN

1. **Risk management system** for AI systems
2. **Governance** and **quality of datasets** used to build AI systems
3. **Record keeping** through **logging capabilities** by AI systems
4. **Transparency** and **information provisions** to the users of AI systems
5. **Human oversight** of AI systems
6. **Accuracy specifications** for AI systems
7. **Robustness specifications** for AI systems
8. **Cybersecurity specifications** for AI systems
9. **Quality management system** for providers of AI systems, including post-market monitoring process
10. **Conformity assessment** for AI systems

Source: European Commission: https://ec.europa.eu/docsroom/documents/52376

Recent Work Items relevant for SR (1/11)

Topic 1: „Risk management for AI systems“

- ISO/IEC 23894 **- Guidance on risk management**
- PWI CEN-CENELEC EN **„AI risk catalogue and risk management“**
- Lack of „**quality management system”** in the AI Act context

Source: JTC 21 Work Programme for the Standardization Request

Recent Work Items relevant for SR (2/11)

Topic 2: „Governance and quality of datasets used to build AI systems"

- ISO/IEC 8183 "Data lifecycle"
- ISO/IEC 5259 "Data quality for analytics and machine learning"
  - p. 2: "Data quality for analytics and ML - **measures**"
  - p. 3: "Data quality for analytics and ML - **management requirements and guidelines**"
  - p. 4: "Data quality for analytics and machine learning (ML) - **data quality process framework**"
- Lack of „**Data quality and bias treatment"** in the AI Act context

Source: JTC 21 Work Programme for the Standardization Request

Recent Work Items relevant for SR (3/11)

Topic 3: „Record keeping through built-in logging capabilities in AI systems“

- ISO/IEC 42001 „**Management system**“
- CEN-CENELEC: **AI system logging** (draft), planned parallel development
- CEN-CENELEC: **AI trustworthiness characterisation** (draft)

Source: JTC 21 Work Programme for the Standardization Request

Recent Work Items relevant for SR (4/11)

Topic 4: „ Transparency and **information provisions** to the users of AI systems**“**

- CEN-CENELEC EN: **AI trustworthiness characterisation** (draft)

- Lack of „Transparency of AI systems” in the AI Act context

  - ➢ complement ISO/IEC 12792 (draft) „Transparency taxonomy of AI systems“

Source: JTC 21 Work Programme for the Standardization Request

Recent Work Items relevant for SR (5/11)

Topic 5: „**Human oversight** of AI systems“

- CEN-CENELEC EN: **AI trustworthiness characterisation** (draft)

Source: JTC 21 Work Programme for the Standardization Request

Recent Work Items relevant for SR (6/11)

Topic 6: „**Accuracy specifications** for AI systems“

- CEN-CENELEC: Accuracy of NLP systems (planned parallel development) (envisaged)
- **Gaps:**
  - Accuracy for computer vision
  - Guidelines for accuracy improvement
  - Guidelines and requirements for accuracy threshold
  - Accuracy of classication systems
    - (maybe adoption of ISO/IEC TS 4213 - Assessment of ML classification performance)
  - Accuracy of AI systems for regression, recommendation and clustering

Source: JTC 21 Work Programme for the Standardization Request

Recent Work Items relevant for SR (7/11)

Topic 7: „**Robustness specifications** for AI systems“

- ISO/IEC 24029-3 „Assessment of the robustness of neural networks“ (draft)

  - Part 3: Methodology for the use of statiscal methods (parallel development planned)

- **Gaps**:
  - Robustness taxonomy for NLP
  - Robustness taxonomy for computer vision
  - Guidelines for robustness improvement
  - Guidelines for robustness threshold definition
  - Robustness assessment for non-neural AI systems (including other ML and symbolic AI)

Source: JTC 21 Work Programme for the Standardization Request

Recent Work Items relevant for SR (8/11)

Topic 8: „**Cybersecurity** specifications for AI systems"

- EN ISO/IEC 27001 "Information security management systems"
- ISO/IEC 27090 "Guidance for addressing security threats and failures in artificial intelligence systems"
- ISO/IEC 27091 "Artificial Intelligence - Privacy protection"
- CEN-CENELEC EN: **AI trustworthiness characterisation** (draft)

Source: JTC 21 Work Programme for the Standardization Request

Recent Work Items relevant for SR (9/11)

Topic 9: „ **Quality management system** for providers of AI systems“

- EN ISO/IEC 27001 “Information security management systems”
- ISO/IEC 42001 „Management system“

Source: JTC 21 Work Programme for the Standardization Request

Recent Work Items relevant for SR (10/11)

Topic 10: „**Conformity assessment** for AI systems“

- ISO/IEC 42001 „Management system“
- ISO/IEC 42006 “Requirements on bodies performing audit and certification of AI management systems”
- **Gaps:**
  - Data quality and bias treatment
  - Testing of AI systems (maybe adoption of ISO/IEC AWI TS 29119-11)
  - Requirements on bodies performing audit and certification of AI systems
  - Competence requirements on AI systems auditors and professionals
  - Conformity assessment framework in the context of the AI Act

Source: JTC 21 Work Programme for the Standardization Request

Recent Work Items relevant for SR (11/11)

ETSI Securing Artificial Intelligence (SAI)

- Traceability of AI models
  - Ownership right protection
  - AI-specific prevention of model misuse
  - ML watermarking
- Manipulation of Multimedia Identity Representations (e.g. fakes)
- Collaborative AI
- Role of Hardware in Security of AI (e.g. „Trusted Execution Environment“)
- Attack types during training and inference & mitigation

Source: JTC 21 Work Programme for the Standardization Request

CEN-CENELEC JTC21 „Artificial Intelligence“

- Topic Group on Cybersecurity (CEN/CLC JTC21, JTC13, ENISA, ETSI SAI, …)
  - Focus: essential requirements on Cybersecurity for AI Act
  - Connect topics „AI systems“, „Cybersecurity“ & „High-risk applications“
  - Identification of Cybersecurity goals (e.g. CIA, traceability, XpI)
  - Overview of threats
  - Gap analysis

Source: JTC 21 Work Programme for the Standardization Request

Bundesnetzagentur

European standardisation synergy: towards the AI Act

Taras Holoyad
AI Standardisation

06131/18 1459
Taras.Holoyad@BNetzA.de