

Forschungsidee: Closed-Loop-Validierung des EU AI Act im Requirements Engineering – Kurzbeschreibung

Glossar

- High-Level Requirements (HLR): regulatorische Pflichten/Schutzziele auf hoher Ebene.
- Low-Level Requirements (LLR): konkrete, testbare Projektanforderungen mit Akzeptanzkriterien.
- Traceability: bidirektionale Nachvollziehbarkeit zwischen HLR und LLR inkl. Audit-Log.
- Human-in-the-Loop (HITL): qualitätssichernde menschliche Reviews und Freigaben.
- Foundation Models (FM): vortrainierte KI-Grundmodelle.
- Large Language Models (LLM): große Sprachmodelle für Textaufgaben.
- Requirements Traceability Matrix (RTM): Zuordnungstabelle (z. B. HLR↔LLR).

Warum jetzt ist der richtige Zeitpunkt

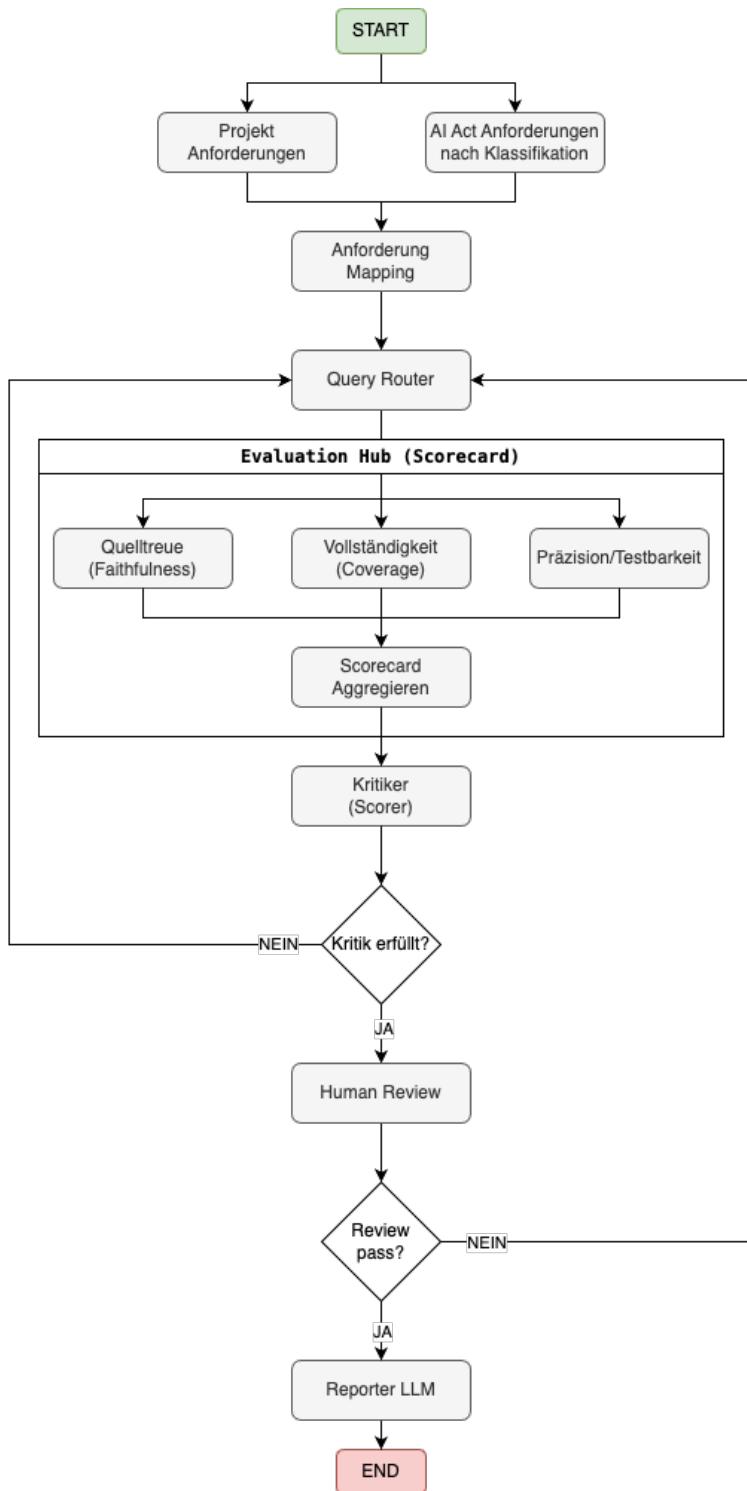
Die EU-Regeln für künstliche Intelligenz (EU AI Act) werden bald umgesetzt, und es entstehen gemeinsame Normen und Richtlinien. Die Teams, die für die Erfüllung dieser Regeln verantwortlich sind, müssen ihre Pflichten in konkrete Anforderungen umwandeln, die kontinuierlich überprüft und dokumentiert werden können. Leider fehlt es heute oft an einem durchgängigen Ansatz für die Überprüfung, die benötigten Daten sind begrenzt, die Messgrößen sind nicht standardisiert und die Überprüfung wird hauptsächlich manuell durchgeführt.

Meine Hypothese und Vision

Ich denke, dass ein agentenbasierter Ansatz, der auf Foundation-Modellen, Conversational-Agents und menschlicher Überprüfung (HITL) basiert, die Erfüllung der EU-AI-Act-Pflichten im Requirements Engineering unterstützen kann. Dieser Ansatz würde es ermöglichen, die Anforderungen transparent und überprüfbar zu machen, die Abdeckung zu erhöhen und den Betrieb auditierbar zu halten.

Die Idee (kurz und konkret)

Ich nutze eine kuratierte EU-AI-Act-Wissensbasis als HLR-Checklisten und Prüfkriterien (ausgerichtet an ISO/IEC/IEEE 29148). Ein Validierungsagent analysiert verschiedene RE Artefakten z.B Spezifikationen, User Stories und RTMs, prüft nach Abdeckung und Konsistenz zwischen HLR und LLR und belegt Befunde mit Quellen. Dabei wird zunächst eine Klassifizierung von Risiken durchgeführt, diese Risikoanforderungen (HLR) werden gegen Projektanforderungen (LLR) gemapp und validiert, um belegbare, testbare Zuordnungen mit hoher Abdeckung zu erhalten. Conversational Agents klären Unschärfen und sammeln Evidenz. Der HITL-Review priorisiert Befunde, setzt Freigaben/Policy-Gates und führt ein Audit-Log. Feedback aus Reviews fließt in Wissensbasis, Prompts und Regeln zurück (Closed-Loop).



Der Fluss zeigt:

Anforderung-Mapping → Query Router → Evaluation Hub (Scorecard: Quelltreue, Vollständigkeit, Präzision/Testbarkeit) → Scorecard-Aggregation → Kritiker/Scorer (Schwellen) → Entscheidungsgate → ggf. Human Review (HITL) → Reporter LLM → Artefakte (Mapping-Liste, Scorecards, Audit-Log, Bericht).

Erwartete Ausgaben

- Mapping-Liste: HLR mit LLR-Subliste und belastbaren Zitationen.

- Scorecards: pro Pflicht und aggregiert (Coverage, Zitat-Treue, semantische Ähnlichkeit, Testbarkeit).
- Review-Protokoll: HITL-Entscheidungen/Freigaben als Audit-Log.
- Bericht/Trace: konsolidierter Validierungsbericht mit Traceability-Verweisen.
- Zusätzlich: Grundlage für synthetische/augmentierte Testdaten und Audits.

Nutzen

- Transparente, nachvollziehbare Zuordnung Risiko ↔ Projekt mit Citations.
- Höhere Abdeckung kritischer Risiken, weniger manuelle Prüfung.
- Auditfähige Dokumentation und belastbare Basis für PMM/Audits und Testdaten-Augmentation.

Kontakt & Teilnahme

Fragebogen (15–20 Min, freiwillig, anonym): [Closed-Loop-Validierung des EU AI Act im RE](#)

Kontakt: emmy.lai@iais.fraunhofer.de

Beispiel Ausgabe:

HLR- ID	AI-Act Bezug	Zitat (sinngemäß)	LLR- ID	LLR (testbar)	Akzeptanzkriterien	Evidenz	Traceability
HLR- 9.1	Art. 9(1)	Anbieter richten ein dokumentiertes Risikomanagementsystem ein und pflegen es.	LLR- 9.1a	Risikoregister in Tool X mit Feldern: Hazard, Harm, Severity, Likelihood, Residual Risk, Mitigation, Owner.	100% Hazards haben vollständige Felder; wöchentlicher Auto-Report ohne Lücken.	Link: RiskRegister v1.3; CI-Report #224	HLR-9.1 ↔ LLR-9.1a (1:n)
HLR- 9.2	Art. 9(2)	Kontinuierlich und iterativ über den gesamten Lebenszyklus.	LLR- 9.2a	CI-Pipeline triggert Risk-Review bei jeder Modell-/Prompthub-Änderung.	≥95% relevanter Commits erzeugen Risk-Review-Job; Job-Log vorhanden.	CI Logs run- 8742; Policy Gate P2	HLR-9.2 ↔ LLR-9.2a