CLARA

A Hybrid LLM-Rule System for EU AI Act Risk Classification

Vasco Alves[1], Iakovina Kindylidi[2,3], André Carreiro[1]

[1]Fraunhofer Portugal AICOS, Porto, Portugal
[2]NOVA School of Law, Lisbon, Portugal
[3]Vieira de Almeida Associados (VdA), Lisbon, Portugal

Read the Paper!

CLARA turns free-form AI system descriptions into **transparent, legally grounded EU AI Act risk classifications** by combining **LLM reasoning** with **expert-authored rules**.

Motivation & Objectives

The **EU AI Act** poses a regulatory challenge, as it is **complex** and **hard to apply** consistently in practice. Existing compliance tools lack **flexibility**, **legal precision**, or **explainability**. Therefore, enabling and improving **automated AI Act risk classification** is a necessity.

Figure 1: AI Act risk levels.

CLARA Overview

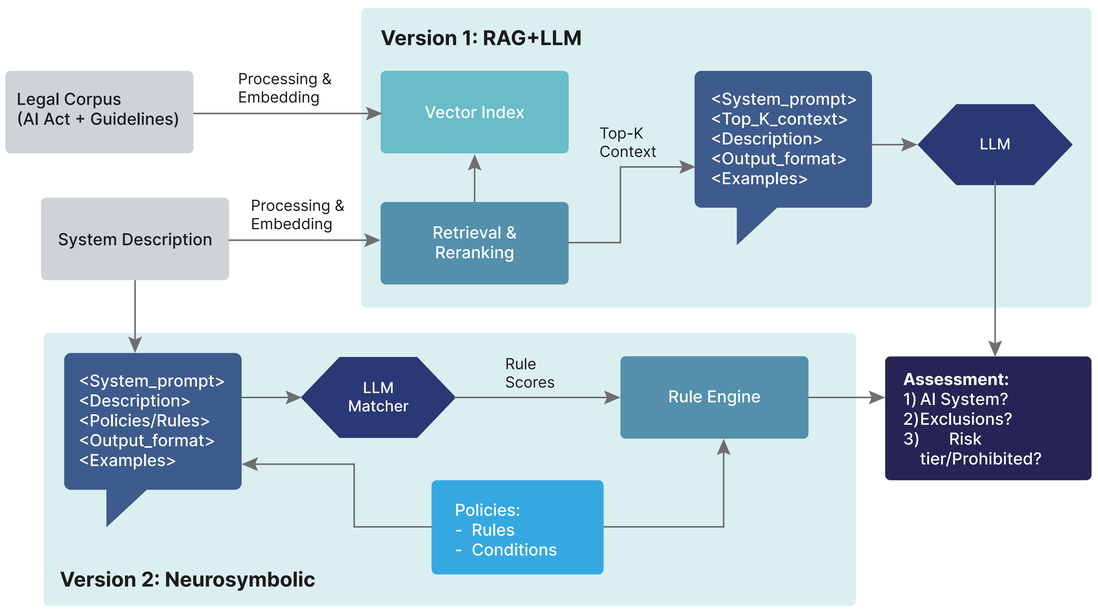

Figure 2: Diagram illustrating CLARA's workflow for the two versions considered: 1) LLM+RAG, 2) Neurosymbolic.

Three-step workflow: AI system qualification → Material scope and exclusions/exceptions → Risk tier suggestion

Two pipelines:

- **RAG+LLM**: end-to-end reasoning over retrieved legal text
- **Neurosymbolic**: LLM condition matching + deterministic rule engine

Evaluation

Four assessments for **16 realistic descriptions**:

- Independent assessment by a **legal expert**
- **ChatGPT** (5.1)
- **RAG+LLM** pipeline (GPT-5-nano)
- **Neurosymbolic** pipeline

Focus on:

- Consistency of reasoning
- Validity of legal references
- Follow-up questions

In the example shown in Fig. 3 regarding a **lifestyle-scoring** system:

- Both **ChatGPT** and **Neurosymbolic** identify it as an **AI system**; **RAG+LLM** returns **Undetermined**.
- **Neurosymbolic** correctly classified it as **probably prohibited**. **ChatGPT** and **RAG+LLM** approaches classified it as **high-risk**.

CLARA - Example

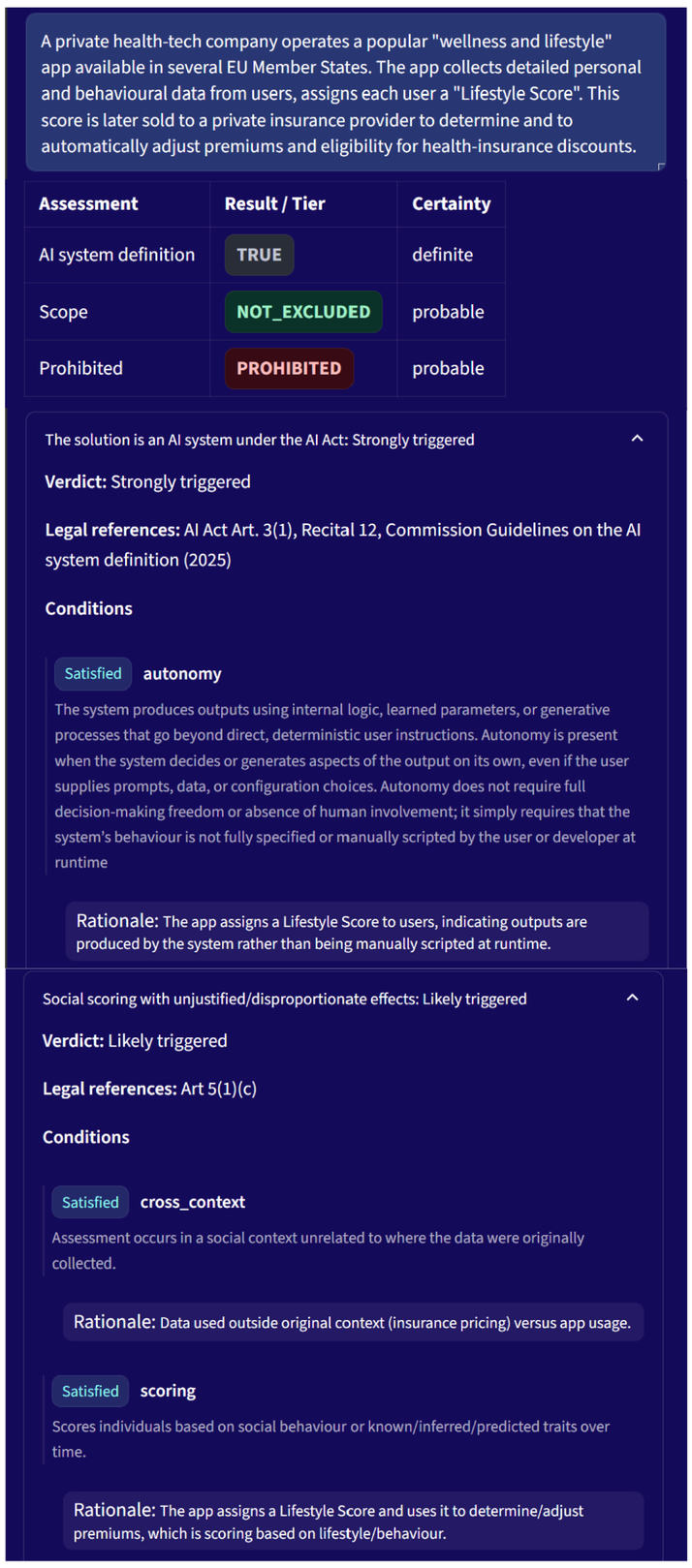

Figure 3: Screenshot of Neurosymbolic version.

Conclusions

CLARA is a tool that combines LLM-based semantic reasoning with deterministic rule logic to support early-stage AI Act self-assessment:

- **Neurosymbolic** reasoning **improves robustness** over LLM-only legal assessment.
- Explicit rules enable **traceable** and **reproducible** risk classification.
- Transparent **explanations** and **follow-up** questions improve understandability

Future work includes expanding legal rule coverage and integrate interactive, human-in-the-loop clarification.

References

Davvetas et al., TAI Scan Tool: A RAG-Based Tool With Minimalistic Input for Trustworthy AI Self-Assessment, arXiv:2507.17514.
Kovari et al., Let's have a chat with the EU AI Act, arXiv:2505.11946.
Baig et al., EU AI Act Applicability Assessment & Compliance Checker, Securiti.
Future of Life Institute, EU AI Act Compliance Checker.

Acknowledgements

This work was supported by European through the Recovery and Resilience Plan, via the project "Center for Responsible AI", with identification number C645008882-00000055.