Course : Generative AI, understanding, justifying and auditing the explicability of LLM models

Explain, document and audit LLM decisions in your AI projects

Practical course - 3d - 21h00 - Ref. LLO

Price : 2010 € E.T.

NEW

This course provides participants with the methods and tools to analyze, explain and audit LLM decisions. You will learn how to document, justify and supervise generative AI according to ethical and regulatory principles.

Teaching objectives

At the end of the training, the participant will be able to:

- Understand the strategic stakes of explicability for the adoption of generative AI.
- Assess the business risks associated with bias, error and opacity of LLMs.
- Design internal policies to guarantee transparency and accountability in AI use.
- Deploy explainability methods and tools to reinforce trust and control.
- Align AI governance with regulatory frameworks and compliance requirements.
- Manage the audit and ongoing supervision of GenAI projects within the company.

Intended audience

Architectes IA, data scientists, PO, juristes, fonctionnels, auditeurs internes et tout professionnel impliqué dans la responsabilité des systèmes IA.

Prerequisites

Basic knowledge of AI or LLMs. Practice in reading and writing prompts.

Course schedule

PARTICIPANTS

Architectes IA, data scientists, PO, juristes, fonctionnels, auditeurs internes et tout professionnel impliqué dans la responsabilité des systèmes IA.

PREREQUISITES

Basic knowledge of AI or LLMs. Practice in reading and writing prompts.

TRAINER QUALIFICATIONS

The experts leading the training are specialists in the covered subjects. They have been approved by our instructional teams for both their professional knowledge and their teaching ability, for each course they teach. They have at least five to ten years of experience in their field and hold (or have held) decision-making positions in companies.

ASSESSMENT TERMS

The trainer evaluates each participant's academic progress throughout the training using multiple choice, scenarios, hands-on work and more.
Participants also complete a placement test before and after the course to measure the skills they've developed.

1 Why explicability is essential

- Definitions: interpretability vs. explicability.
- Specific features of generative models and LLMs.
- Critical use cases: legal, medical, HR, finance.
- Risks associated with lack of explanation: trust, adoption, compliance.

2 Explicability of LLMs, limits and levers

- How do LLMs work: black boxes or reducible systems?
- Prompts, memory, outputs: where are the biases?
- Structural limitations: instability, hallucination, lack of traceability.
- Reproducibility: challenge or mirage?

3 Explainability methods in the GenAI context

- Interpretability-oriented prompt engineering.
- Chain-of-thought" approaches, step-by-step reasoning.
- Generated justification vs. calculated proof.
- Elements observable in LangChain: logs, agents, tools.

Hands-on work

LLM response analysis and step-by-step justification. Comparative analysis of correct / incorrect generation. Reconstruction of the reasoning chain. Visualization of context and complete prompt

4 Trace, understand and explain via LangChain

- Traceable components in LangChain: agents, tools, chain logs.
- Logging, callback handlers, explicit prompt templates.
- Introduction to TruLens, PromptLayer, Helicone, LangSmith.
- Creation of a traceable pipeline.

5 Using ontologies and graphs to explain

- Structure knowledge to explain it better.
- Knowledge graphs + LLM = interpretable context.
- Business ontologies: intelligible explanations for the user.
- Dialogue between LLM agent and structured graph.

Hands-on work

Creation of an explainable assistant. Agent justifies its answers using a graph/ontology. Self-explanatory prompts. Complete logging from request to response.

6 Explicability and regulatory framework

- What the RGPD and the IA Act demand (right to explanation, transparency).
- Requirement for documentation, logs, reproducibility.
- Explainable interfaces: how to display an intelligible rationale.
- The role of explicability in DPIAs and risk assessments.

TEACHING AIDS AND TECHNICAL RESOURCES

• The main teaching aids and instructional methods used in the training are audiovisual aids, documentation and course material, hands-on application exercises and corrected exercises for practical training courses, case studies and coverage of real cases for training seminars.
• At the end of each course or seminar, ORSYS provides participants with a course evaluation questionnaire that is analysed by our instructional teams.
• A check-in sheet for each half-day of attendance is provided at the end of the training, along with a course completion certificate if the trainee attended the entire session.

TERMS AND DEADLINES

Registration must be completed 24 hours before the start of the training.

ACCESSIBILITY FOR PEOPLE WITH DISABILITIES

Do you need special accessibility accommodations? Contact Mrs. Fosse, Disability Manager, at psh-accueil@orsys.fr to review your request and its feasibility.

7 GenAI system audit methods

- Create an explanatory log: prompt + context + sources + reasoning.
- Quality control of generated responses (hallucination, consistency, bias).
- Inclusion of "critical" agents or explanation scores.
- Human evaluation of reasoning.

Hands-on work

Build an auditable response system. Case study: legal or HR assistant. Set up a complete flow with justification, audit log. Demo of an explanatory interface (textual + graphic).

Dates and locations

REMOTE CLASS
2026 : 29 June, 28 Sep., 7 Dec.

PARIS LA DÉFENSE
2026 : 22 June, 21 Sep., 30 Nov.