

# Guilhem Fouilhé

## PhD Candidate in Computer Science & AI

✉ g.g.a.fouilhe@gmail.com

☎ +33 6 46 49 44 93

📍 Toulouse, France

🌐 gfouilhe.github.io

🔗 gfouilhe



## Profile

PhD candidate in Computer Science working on explainability and interpretability of Large Language Models (LLMs), Agentic Systems, and AI planning. Strong background in Mathematics and Machine Learning (ML), with hands-on experience building full-stack AI applications. Open to research or applied AI roles from late 2026.

## Education

- 2023 – present **Ph.D. in Computer Science**, Université de Toulouse (IRIT / ANITI).  
Advisors: Nicholas Asher & Philippe Muller. Expected defense: December 2026.
- 2021 – 2023 **M.Sc. in Computer Science & Mathematics for AI**, Université Toulouse III – Paul Sabatier.  
*Valedictorian (major de promotion), with honours.*
- 2017 – 2021 **B.Sc. in Mathematics**, Université de Montpellier.  
First two years as a Mathematics–Physics double major. With honours.

## Research Experience

- 2023 – present **Doctoral Researcher, IRIT & ANITI, Université de Toulouse**  
Explaining complex AI models interactively and approximately: explainability and interpretability of large language models and automated planning. Member of the EU project TUPLES.
- 2023 (6 mo.) **Research Intern, IRT Saint Exupéry**  
Supervisors : David Vigouroux & Agustin Picard.  
*Influence functions for global explainability of deep neural networks.*
- 2022 (3 mo.) **Research Intern, CerCo – CNRS** (funded by ANITI).  
Supervisors : Grégory Faye & Rufin VanRullen.  
*Mathematical models for predictive coding in vision.* Led to [5].
- 2021 (2 mo.) **Research Intern, IMT – Institut de Mathématiques de Toulouse** (funded by CIMI).  
Supervisors : Grégory Faye & Christophe Besse.  
*Sharp thresholds for bistable reaction–diffusion equations.* Led to [6].

## Publications

- 1 **Fouilhé, G.**, Eifler, R., Poché, A., Thiébaux, S., & Asher, N. (2026). Exploring Plan Space through Conversation: An Agentic Framework for LLM-Mediated Explanations in Planning. *arXiv*. Under review. arXiv: 2603.02070
- 2 Naim, O., **Fouilhé, G.**, & Asher, N. (2025, September). Re-examining Learning Linear Functions in Context. In *KI 2025: Advances in Artificial Intelligence* (pp. 104–117). doi:10.1007/978-3-032-02813-6\_8
- 3 Eifler, R., & **Fouilhé, G.** (2025). Ipxeco: a platform for iterative planning with interactive goal-conflict explanations. *Late-breaking work, Demos and Doctoral Consortium, colocated with The 3d World Conference on eXplainable Artificial Intelligence*.
- 4 **Fouilhé, G.**, Eifler, R., Thiébaux, S., & Asher, N. (2025). Conversational goal-conflict explanations in planning via multi-agent llms. *Workshop on Planning in the Era of LLMs (LM4Plan @ AAAI 2025)*.
- 5 Faye, G., **Fouilhé, G.**, & VanRullen, R. (2023). Mathematical derivation of wave propagation properties in hierarchical neural networks with predictive coding feedback dynamics. *Bulletin of Mathematical Biology*, 85(9), 80.

- 6 Besse, C., Capel, A., Faye, G., & **Fouilhé, G.** (2022). Asymptotic behavior of nonlocal bistable reaction-diffusion equations. *Discrete and Continuous Dynamical Systems - B*, 0-0.  
doi:10.3934/dcdsb.2022211

## Teaching

2023 – 2025	Lectures, tutorials and lab sessions for undergraduate (L1–L3) and graduate (M1) students at Université Toulouse III – Paul Sabatier and the UPSSITECH engineering school.
Computer Science	Advanced Algorithmics, Algorithms & Data Structures, Databases, Object-Oriented and Advanced Programming, Artificial Intelligence (Machine & Deep Learning).
Mathematics	Functions & Calculus, Set Theory.
Outreach	Co-authored a national online module on the use of generative AI for students (Université Numérique).

## Projects & Software

2026	<b>SSA-Luciole-1B</b> – Pretraining a billion parameters state-of-the-art LLM from scratch. HPC GPU Training and Evaluation (Pytorch Lightning, Nvidia NeMo, Hugging Face, Custom Triton Kernel).
2025	<b>IPEXCO</b> – platform for iterative planning with interactive goal-conflict explanations. Full-stack development (Angular, Node.js / Express, MongoDB, OpenAI API). <b>TempEval Visualization</b> – interactive visualization for temporally annotated data (TimeML / TimeBank / TempEval).
2022	<b>Linguistic markers of dementia</b> – academic project, IRIT (supervised by Chloé Braud): literature review and predictive modelling.
2021	<b>Pl@ntNet</b> – visualization tools for a Pl@ntNet dataset, Institut Montpellierain Alexandre Grothendieck.

## Skills

Programming	Python (primary), C/C++, Java, JavaScript/TypeScript; notions of Rust, R, Julia, Matlab, OCaml, SQL.
ML & Data Science	<b>PyTorch / Lightning / NeMo / Triton, Hugging Face</b> (Transformers, Datasets, PEFT, TRL), lm-evaluation-harness, <b>TensorFlow/Keras</b> , scikit-learn, NumPy/SciPy, Pandas, spaCy/NLTK, OpenAI API, VLLM, Gradio.
Web & Tooling	Angular, FastAPI, Node.js, MongoDB; Git, Docker, Linux, $\LaTeX$ .
Languages	French (native), English (fluent, C2).

## Awards & Funding

2021 – 2023	<b>ANITI Graduate School Fellowship</b> (Master's, awarded in both M1 and M2).
2023	<b>Valedictorian</b> of the M.Sc. in Computer Science & Mathematics for AI.
2021 / 2022	<b>Competitive research internship grants</b> from CIMI and ANITI.

## Academic Service & Leadership

2024 – present	<b>PhD &amp; Postdoc Representative, ANITI</b> Organize seminars, workshops and social events for the AI cluster's early-career researchers.
2025	<b>Scientific Evaluator</b> , Airbus Beluga AI Challenge – Explainability Track.
2024 – 2025	Invited talks (LAAS-CNRS, IRT Saint Exupéry); LxMLS 2024 summer school (Lisbon); TUPLES project meetings across Europe (Prague, Saarbrücken, Leuven).
2018 – 2021	<b>Student governance (Montpellier)</b> . President of the student union (AGEM, 2019–2020): led a 15-person team and a federation of 75,000+ students; elected to university, faculty and CROUS councils.