

Maël Jullien

Cambridge, UK

mael.jullien@gmail.com | mael7307.github.io | [LinkedIn](#) | [GitHub](#) | [Google Scholar](#) | [ORCID](#)

AI R&D Engineer: Specialising in agentic AI systems for expert legal and medical natural language tasks. End-to-end experience researching, prototyping, developing, and optimising production LLM applications for automation & augmentation. PhD in Computer Science published at ACL, EMNLP, NAACL, and SemEval.

Experience

Research and Development Engineer | Luminance Feb 2026 – Present
Cambridge, UK

- Led the research and development of **Auto-Markup**, decomposing contract analysis, editing, and verification across stateful agents with parallel tool execution and specialist sub-agent judges. Reduced end-to-end latency by **50%** while improving benchmark performance and recall.
- Led the development and optimisation of **Auto-Templating**, an agent combining autonomous editing, structural validation, repair, and simulated-user evaluation. Reduced a multi-hour workflow to minutes.
- Architected an **evaluation system for expert-level contract markups**, combining deterministic scorers, annotations and specialised LLM judges to diagnose regressions and guide architectural changes.
- Developed an **Agents-as-a-User** framework for the workflow platform, producing reports, tagging users, advancing workflow stages, and sending notifications through an extensible agent framework and tool suite.

Research Intern | Idiap Research Institute 2025 – 2026
Martigny, Switzerland

Education

PhD in Computer Science | University of Manchester 2020 – 2026
Thesis: “Evaluating and Controlling Natural Language Inference over Clinical Trial Texts”

MSc in Computer Science and Artificial Intelligence; Distinction | University of Nottingham 2020
Thesis: “Histological Image Registration using Deep Convolutional Features”

BSc Mathematics; First Class | University of Leicester 2019
Thesis: “Convolutional Neural Networks for Audio Classification”

Skills

- **Agentic AI:** Multi-agent architectures, stateful tool-using agents, parallel agent orchestration, model and skill routing, closed-loop autonomous workflows, MCPs, simulated-user agents and long-horizon workflow execution.
- **LLM Evaluation:** Agent evaluation, dataset construction, LLM-as-a-judge, expert-grounded benchmarks, deterministic and model-based scoring.
- **Machine Learning:** Text Classification, Natural Language Inference, Question Answering, Retrieval-Augmented Generation (RAG), Hyperparameter Tuning, Information extraction, Prompt engineering, Document-level language understanding, Models from 110M to 1T parameters.
- **Computing:** Python, PostgreSQL, Codex/Cursor/Claude-code agentic software engineering, GitLab CI/CD, GitHub

Selected Publications

SemEval-2023 Task 7: Multi-Evidence Natural Language Inference for Clinical Trial Data | Jullien, M., Valentino, M., Frost, H., O'Regan, & Freitas, A. SemEval 2023 @ ACL 2023

NLI4CT: Multi-Evidence Natural Language Inference for Clinical Trial Reports | Jullien, M., Valentino, M., Frost, H., O'Regan, P., Landers, D., & Freitas, A. EMNLP 2023

SemEval-2024 Task 2: Safe Biomedical Natural Language Inference for Clinical Trials | Jullien, M., Valentino, M., & Freitas, A. SemEval 2024 @ NAACL 2024

Compartmentalised Agentic Reasoning for Clinical NLI | Jullien, M., Xu, L., Valentino, M., & Freitas, A. Findings of ACL 2026

Dissecting Clinical Reasoning in Natural Language Inference for Large Language Models | Jullien, M., Valentino, M., Ranaldi, L., & Freitas, A. Findings of ACL 2026