

Da Silva Gameiro Henrique

EPFL Data Science engineer | Experience in Research | Interested in beneficial, robust AI

✉ henrique.dasilvagameiro@gmail.com ☎ +41 78 711 36 71 📍 Collombey, Valais Switzerland 🇨🇭 Swiss

🐙 GitHub 🔗 LinkedIn

PROFESSIONAL EXPERIENCE

Research Trainee (6 months)

2025 – 2025

RIKEN AIP

Tokyo

- **Task:** Compare and scale existing techniques to link **LLM prediction** to training data and for the **analysis of AI models at scale**.
- **Work done:** Implemented current best methods, created experiments to compare state-of-the-art methods, developed a **innovative method**. Improvement of Generative AI **transparency**.
- **Output:** Work in progress scientific paper and library.
- **Skills learned:** GPU clusters, Reproducing paper results, dealing with **large-scale** constraint in methods to create **scalable** experiments using **profiling** and performance debugging, **ML research** development lifecycle in a large scientific institution.

Master Intern (6 months)

2024 – 2024

Cyber-Defence Campus Lausanne

Lausanne

- **Task:** Evaluate **LLM detectors** against attacks and develop watermarking defense for **Responsible AI**.
- **Work done:** Developed adversarial testing of LLM detectors, evaluation of adversarial attacks, created a LLM detector **benchmark**, developed watermarking for text as a defense, Model **inference and evaluation** with and without watermark.
- **Output:** Submitted **Scientific paper** to **NLP conferences**, book chapter into the open-access "LLM in Cybersecurity book" edited by Springer together with other experts of the field.
- **Skills learned:** Self-start a research project independently with few supervision, **safety thinking**, **problem-solving** in scientific environment, robust scientific methodology, LLM evaluation, benchmarking pipelines of commonly used LLMs such as **Llama** and **Mistral**.

SKILLS

Technical skills

- **Programming languages:** Python, SQL, Scala, Java, Kotlin, C, C#, C++
- **ML frameworks:** PyTorch, NumPy, Hugging Face, Distributed AI Training (Accelerate), Accuracy evaluation frameworks, Training and evaluation logging (W&B), Retrieval Augmented Generation (LangChain, PydanticAI)
- **Data & Systems:** Pandas, Spark, Kafka, Hive, Docker, Singularity, Linux, CI/CD, Git, Visualization (Matplotlib, Seaborn)

Soft skills

- **Professional:** Adversarial and critical thinking, Open-source, Independent, Pragmatic (focus on most impact and deliver), Navigating complex codebases.
- **Qualities:** Curiosity, Ambitious, Perseverance, Team player, Autonomous and Self-starter, Long-term vision.
- **Hobbies:** Video games, History, Philosophy, Learning languages, Running, Bouldering, Boxing
- **Natural languages:** French (native), English (C1 Cambridge), Portuguese, German, Japanese

SELECTED PROJECTS

GPT model detectability (Individual project with an EPFL lab)

2022 – 2022

- **Task:** Detect AI-generated text.
- **Work done:** Designed and tested adversarial training of **GPT-2**, Analysis and defense against harmful outputs.
- **Output:** Submitted paper to NLP conference.

- **Skills learned:** Machine learning experimentation, Transformer-based models design, **Generative AI**, open-goal projects, scientific writing, ML frameworks **PyTorch**, **Hugging Face**, working with AI Architecture development, **AI training** and LLM (Large Language Model) **safety**.

Conversational LLM Educator with SFT/RLHF Training (Team of 3)

2023 – 2023

- **Task:** Build conversational LLM for education.
- **Work done:** LLM fine-tuned with GPT-3.5 output with Chain-of-Thought reasoning (**distillation**), RLHF **alignment** using Reddit data, Multi-round answers.
- **Skills learned:** Dataset search, prompt engineering, reinforcement learning for alignment, parts of MLOps technical stack (Machine Learning Operations): **Monitoring & observability** (W&B), Training and evaluation of large-scale neural networks (Fine-tuning) and **Post-training evaluation**.

Large beer review dataset analysis (Team of 4)

2022 – 2022

- **Task:** Develop a data story using a raw **large-scale dataset**, produce data story website.
- **Work done:** Constructed a **full data pipeline** from raw dataset, **data curation** to insights, analyzed consumer categories using **clustering** and sentiment analysis (BERT).
- **Output:** Created a website showing the outcome of the data analysis.
- **Skills learned:** Teamwork, Large-scale data processing, visualization with interactive plots, analytical skills.

Android blind test game app (Team project)

2022 – 2022

- **Task:** **Agile** development of android app in a team of 6 using **Scrum**, ensure at least 80% code coverage to emulate production-level quality.
- **Work done:** Weekly shipping of features with sprints. Created a **database** to log user's activity and create a user profile menu as well as a leaderboard. Helped with game loop algorithm, online features, voice recognition and testing.
- **Output:** Working android app game that ended 2nd out of 20 in the internal competition by vote.
- **Skills learned:** Teamwork with Scrum, Maintain **DevOps stack** (**CI/CD**, **Test design** and **debugging**, **version control systems**, **deployment**), Stable and scalable software with complex software design, extensive documentation.

EDUCATION

Data Science Master's degree

EPFL
(done)

2022 – 2025
Lausanne, Switzerland

Computer Science Bachelor's degree

EPFL

2019 – 2022
Lausanne, Switzerland

Latin-English & Physics maturité

Lycée-collège de l'Abbaye de Saint-Maurice

2014 – 2019
Saint-Maurice,
Switzerland