# STEFANO VIEL

viel.stefano01@gmail.com ◇ +39 377 4497798

[GitHub](#) ◇ [LinkedIn](#) ◇ [Portfolio](#)

## EDUCATION

**MSc in Data Science**, EPFL                                        September 2023 - Expected July 2025

- Courses: Data analysis, Machine learning, Causal inference, Markov chains, Reinforcement learning, Modern natural language processing, Network machine learning, Large scale data science.
- **GPA:** 5.5/6

**BSc in Computer Science**, University of Trento                    September 2020 - July 2023

- Relevant courses: Programming languages (Python, C++, Java), Algorithms and data structures, Databases.
- **GPA:** 28.57/30 (top 5% of my year), **Final Grade:** 110/110 with honors.

## WORK EXPERIENCE

**Research Intern**            June 2024 - September 2024 **ETH Zurich**            *Zurich, Switzerland*

- Developed a GPU-accelerated data processing pipeline to categorize 1.7 billion Reddit posts into topics.
- Conducted comprehensive literature review and implemented state-of-the-art topic modeling leveraging BERT embeddings, UMAP, and HDBSCAN for optimal results
- Developed and launched [redditopics.xyz](#), an interactive platform for exploring and accessing the processed dataset
- Technologies: Python, Git, Docker, SQL, cuml (GPU computing), Plotly, Flask, React.

**Research Intern**            March 2023 - June 2023 **Fondazione Bruno Kessler**            *Trento, Italy*

- Developed Python scripts to map Mastodon's decentralized network, scraping over 100,000 decentralized servers and gathering data on 6 million users to map the complete network structure.
- Implemented asynchronous processing and efficient queuing systems to handle API rate limits, storing and analyzing large-scale social network data in MongoDB

## PROJECTS

**Course Assistant Chat-bot for STEM Subjects**                    Modern NLP @ EPFL

- Developed STEMERALD, a STEM course assistant using the Gemma-2b language model, implementing Supervised Fine-Tuning and Direct Preference Optimization.
- Achieved 70% accuracy on multiple-choice benchmarks and reduced the model's memory footprint from 10GB to 2GB for consumer-grade deployment.
- Technologies: Python, Pytorch, Git, HuggingFace, SFT, RLHF.

[Questudying.com](#)                                                Personal project

- Developed an AI-powered personal learning assistant that helps users revise more efficiently by answering questions about the material they upload.
- Questions are tuned based on the previous user's answers to reach the right level of difficulty.
- Technologies: Retrieval-Augmented Generation, Python, Flask, React, Git, Docker, Prompt Engineering.

## CERTIFICATES, AWARDS, AND HOBBIES

**2024** Selected as a mentee for LeadTheFuture, a top STEM mentorship program with an acceptance rate below 15%.

**2024** Selected to participate in Hack the North, winning first prize in the Groq-sponsored category. [Demo Video](#)

**2023** IELTS Certification: Overall Score of 8.0, Equivalent to C1 on the CEFR scale.

**2022** Participated in the Utrecht Debating Society for a duration of six months.