

Nicolò De Sabbata

☎ (+39) 3884756535 | ✉ cndesabbata@outlook.com | 🏠 nicolodesabbata.com | 📷 cndesabbata | 📺 cndesabbata | he/him

Passionate about engineering intelligence. Dedicated to using my skills and knowledge to solve complex problems and create meaningful solutions.

Experience

EPFL - NLP Lab

ML RESEARCHER - INTERN

- Conducting research on reasoning abilities of **Large Language Models**

Lausanne, Switzerland

Oct 2024-Present

AXA

APPLIED SCIENTIST - INTERN

- Responsible for designing, building, and fine-tuning **LLM Agents** up to 70B in size to automate business processes.

Lausanne, Switzerland

Sep 2023 - Feb 2024

Amazon

DATA ENGINEER - INTERN

- Designed and developed a library to automate the monitoring of **ETL pipelines** and **Data Management Systems**
- This tool allows engineers to efficiently **collect metrics**, proactively **predict failures** and **deploy alarms** using AWS CDK

Luxembourg, Luxembourg

Jul 2022 - Oct 2022

Education

Princeton University

VISITING STUDENT RESEARCHER

- Conducted my **Master's Research Project** within the Computational Cognitive Science laboratory of Professor Tom Griffiths.

Princeton, NJ, United States

Mar 2024 - Aug 2024

EPFL - Swiss Federal Institute of Technology

M.S. IN COMPUTER SCIENCE

- Specializing in AI&ML, student researcher in the NLPLab, GPA 5.3/6

Lausanne, Switzerland

Sep 2021 - Sep 2024

Politecnico di Milano

B.S. IN COMPUTER ENGINEERING

- Graduated with **108/110 (top 5%)**, obtained **Best Freshmen Prize** based on outstanding academic performance.

Milan, Italy

Sep 2018 - Jul 2021

Publications

Rational Metareasoning for Large Language Models

Nicolò De Sabbata, Theodore Sumers, Thomas L. Griffiths. *Published: S2RAS & BML Workshops, NeurIPS 2024*

- Developed a novel method to optimize reasoning in large language models using a novel reward function based on the Value of Computation, enabling models to selectively use intermediate reasoning steps only when necessary.
- Demonstrated the effectiveness of this approach across diverse datasets, reducing inference costs by up to 37% without compromising performance.

Understanding the Limits of Vision Language Models Through the Lens of the Binding Problem

Declan I. Campbell, Sunayana Rane, Tyler Giallanza, Nicolò De Sabbata, Kia Ghods, Amogh Joshi, Alexander Ku, Jonathan D. Cohen, Thomas L. Griffiths, Taylor W. Webb. *Published: NeurIPS 2024*

- Investigated vision-language models' limitations in multi-object reasoning, linking them to cognitive science's binding problem and trade-offs between flexibility and capacity.
- Demonstrated that VLMs face human-like limits in multi-object scene processing in tasks like counting, localization, and visual analogy.

Relevant Projects

AI-Powered Educational Chatbot

- Collected supervised fine-tuning demonstrations by distilling a 100B-parameter scale LLM, and used it to fine-tune GPT2
- Trained a reward model to perform reinforcement learning with human feedback (RLHF) to further improve the quality of generation.

Distributed Movie Recommendation System

- Developed a Spark based distributed recommendation system on millions of reviews
- Implemented a distributed K-Nearest Neighbors algorithm to improve recommendation precision

Noise2Noise Deep learning network

- Implemented a deep convolutional encoder-decoder neural network for image denoising in PyTorch, trained without a clean reference image

Skills

Coding Languages

Python, Scala, Java, Typescript, Javascript, C, SQL, HTML, CSS

Frameworks & Tools

Pytorch, Spark, AWS, Git, Docker, Hadoop, Scikit-learn, TensorFlow, Matplotlib, Seaborn