

# Davide Baldelli

✉ [davide.baldelli@mila.quebec](mailto:davide.baldelli@mila.quebec) | 🏠 [dundalia.github.io](https://dundalia.github.io) | 🐙 [github](https://github.com) | [in linkedin](https://www.linkedin.com/in/davidebaldelli)

## Education

---

### MILA and Polytechnique Montréal

PhD in Computer Engineering

*Montréal, Canada*

*Jan 2025 - Current*

- GPA: 4.0
- **Supervisors:** Sarath Chandar, Amal Zouaq.
- **Courses:** Representation Learning, Reinforcement Learning.

### University of Bologna

MSc in Artificial Intelligence

*Bologna, Italy*

*Sept 2021 - Oct 2023*

- Grade: 110/110 cum laude.
- **Courses:** Programming for Data Science, Knowledge Representation and Reasoning, Statistical Theory and Methods, Machine Learning, Deep Learning, Computer Vision, Natural Language Processing.

### University of Florence

BSc in Mathematics

*Florence, Italy*

*Sept 2017 - Mar 2021*

- Grade: 110/110.
- **Courses:** Mathematical Analysis, Linear Algebra, Group Theory, Geometry, Probability and Statistics, Logic and Computability.

## Research Experience

---

### National Institute of Informatics

MSc Thesis - International Internship Program

*Tokyo, Japan*

*Mar 2023 - Aug 2023*

- **Supervisors:** Akiko Aizawa (National Institute of Informatics), Paolo Torroni (University of Bologna)
- **Title:** A TWO-step LLM Augmented distillation method for passage Reranking

### University of Florence

BSc Thesis

*Florence, Italy*

*Sept 2020 - Mar 2021*

- **Supervisor:** Francesca Romana Nardi (University of Florence)
- **Title** (translated): Big Deviations for Markov Chains

## Publications

---

### PEER-REVIEWED CONFERENCE PROCEEDINGS

**Baldelli, Davide**, Junfeng Jiang, Akiko Aizawa, and Paolo Torroni (2024). “TWOLAR: A TWO-Step LLM-Augmented Distillation Method for Passage Reranking”. *Advances in Information Retrieval*. Vol. 14608. Lecture Notes in Computer Science. Springer, pp. 465–481. DOI: 10.1007/978-3-031-56027-9\_29. URL: [https://doi.org/10.1007/978-3-031-56027-9\\_29](https://doi.org/10.1007/978-3-031-56027-9_29).

### PRE-PRINTS AND WORKING PAPERS

Govindarajan, Prashant, **Baldelli, Davide**, Jay Pathak, Quentin Fournier, and Sarath Chandar (2025). *CADmium: Fine-Tuning Code Language Models for Text-Driven Sequential CAD Design*. arXiv: 2507.09792 [cs.LG]. URL: <https://arxiv.org/abs/2507.09792>.

## Conferences and Presentations

---

### 46th European Conference on Information Retrieval (ECIR 2024)

Oral Presentation

*Glasgow, UK*

*Mar 2024*

- **Paper:** TWOLAR: A TWO-step LLM-Augmented Distillation Method for Passage Reranking

## Professional Experience

---

### MILA

Teaching Assistant - Summer High School Internships

*Montréal, Canada*

*Summer 2025*

- Designed and delivered educational workshops on Machine Learning, Probability, and Statistics for high school students.
- Mentored students in hands-on learning activities, fostering interest in AI research.

### Independent Consultant (Upwork Platform)

*Remote*

Freelance Machine Learning Engineer

*June 2024 - Jan 2025*

- Delivered a variety of machine learning projects for international clients, consistently earning 5-star ratings for quality and speed.

### Loop AI Group

*Full Time - Fully Remote*

Machine Learning Engineer and Master's Thesis Co-Supervisor

*Sept 2023 - Dec 2024*

- Co-supervised a master's thesis titled "Self Assessment Tool for Medical Exam Candidates" in collaboration with the University of Padova. Provided mentorship throughout the entire research process, from problem formulation to final defense. Guided the candidate in literature review, experimental design, machine learning model selection, and evaluation methodology.
- Conducted deep learning project for musical Deepfake detection achieving 94% of F1 score on a proprietary deepfake dataset.
- Implemented BERT-based email classifier tool for an Italian University's administrative office, achieving 98% accuracy.

### Loop AI Group

*Part Time - Fully Remote*

Data Scientist Consultant

*Sept 2022 - Sept 2023*

- Implemented closed-domain generative QA systems on technical topics, exploiting Pretrained Seq2Seq Language Models.
- Implemented rule-based and transformers-based text retrievers.

## Leadership and Community Involvement

---

### Belle Parole Arci APS

*Florence, Italy*

Leadership in Non-Profit Organization

*Nov 2023 - Present*

- Involved in founding and leading Belle Parole Arci APS, a non-profit organization.
- Responsible for organizing community events and managing bar operations as part of the organization's activities.
- Built the website of the organization: [belleparoleaps.it](https://belleparoleaps.it).

## Other Achievements

---

### AWARDS AND RECOGNITIONS

*Jun 2025*      **Selected Participant**, Y Combinator AI Startup School

*San Francisco, CA*

*Jun 2022*      **4th place**, Loop Q Prize Competition

*Italy*

### CERTIFICATIONS

*Apr 2025*      **Agents Course**, Hugging Face

*Remote*

## Skills

---

<b>Programming</b>	Python (fluent); HTML, CSS, R, MatLab, Minizinc, NetLogo (familiar).
<b>Data Science</b>	Numpy, Pandas, Scikit-learn, Matplotlib, Seaborn, Plotly
<b>Deep Learning</b>	Pytorch, Pytorch-lightning, Tensorflow.
<b>NLP</b>	HuggingFace, Transformers, Prompt Engineer, OpenAI GPT.
<b>Miscellaneous</b>	Linux, Shell, L <sup>A</sup> T <sub>E</sub> X(Overleaf/R Markdown), Git, Docker, Figma, Blender.

## Languages

---

<b>English</b>	Professional proficiency
<b>Italian</b>	Native proficiency

# Projects

---

## **Full-Stack AI Agent Platform with LangGraph Integration**

*Montreal, Canada*

Personal Project

*Jul 2025*

- Built a comprehensive full-stack platform for creating and deploying custom AI agents (like DeepResearch) using modern web technologies. Frontend developed with Next.js, React, and TypeScript. Backend implemented with FastAPI and LangGraph framework, supporting multiple AI models.

## **Choosr – Career Orientation Platform for Italian Students**

*Remote*

Creator and Lead Developer

*Sept 2024 – Present*

- Designed and implemented the entire technical stack for a career guidance platform helping Italian high school students navigate university choices.
- Built and integrated comprehensive databases of career paths and Italian university programs.
- Implemented the RIASEC psychometric test to provide personalized career recommendations.

## **Deepdreaming with newest vision architectures**

*Tokyo, Japan*

University of Bologna

*Aug 2023*

- Updated a github repository to support the newest vision models and pretrained checkpoints. With our implementation, images can be deepdreamed either by maximizing the activation of a specific layer or by maximizing the similarity score with a text embedding, allowing the user to inject a visual concept in the form of a phrase while deepdreaming (only with CLIP models).

**References available upon request.**