

Nelly Barret

Post-doctoral researcher at Politecnico di Milano

🌐 <https://nelly-barret.github.io> — [nelly-barret](#) — [ORCID](#)

Research experience

Post-doc: Better real-world health-data distributed analytics research Apr. 2024 - Apr. 2025
Politecnico di Milano *Milan (IT)*

In this European project, we (*i*) propose a **data model** for integrating clinical, genomic, and imaging data; (*ii*) propose and implement an **ETL pipeline** to build a network of interoperable databases. Afterwards, clinicians will **run federated analyses** to better understand 3 genetic rare diseases.

PhD: user-oriented exploration of semi-structured datasets Jan. 2021 - March 2024
Institut Polytechnique de Paris and Inria Saclay *Palaiseau (FR)*

We propose novel **graph-based** methods to apprehend semi-structured datasets (XML, JSON, CSV, RDF, PG) with (*i*) **Entity-Relationship-like summaries** and (*ii*) **entity-to-entity path enumeration and ranking**, both computed from one or several heterogeneous datasets.

Internship: predicting the environment of a neighbourhood with Predihood Feb. - July 2020
LIRIS *Villeurbanne (FR)*

In this interdisciplinary project, we first (*i*) **select relevant features** using a dedicated algorithm and heuristics; (*ii*) **predict** the environment of French neighbourhoods using well-know prediction algorithms; and (*iii*) develop a **framework to run and compare** prediction algorithms.

Internship: comparing and recommending neighbourhoods with VizLiris May - July 2018
LIRIS *Villeurbanne (FR)*

To facilitate the comparison and recommendation of neighbourhoods in France, we (*i*) **integrate cartographic and INSEE** data; (*ii*) **extract high-level indicators** from INSEE data; (*iii*) develop a Web interface for both **predicting similar neighbourhoods and clustering them**.

Scientific publications and softwares

All my publications are available on [my ORCID](#) and they sum to **17 publications**:

- 3 journals (2 Q1¹, 1 Q4, 1 JOSS)
- 5 international publications (2 rank A, 1 rank B, 2 rank C)
- 2 national publications at BDA²
- 2 international workshops (2 rank A*)
- 3 demonstrations (1 rank A, 1 rank B, 1 national)
- 1 thesis manuscript

They led to **7 research softwares**: [BETTER](#) (restricted access), [Abstra](#) (open-source), [PathWays](#) (open-source), [ConnectionStudio](#) (open-source), [Predihood](#) (open-source), [GeoAlign](#) (open-source), and [VizLIRIS](#) (open-source).

¹One of them is under review, thus not yet shown on my ORCID record.

²French-speaking national data management conference.

Education

PhD degree, Computer science track

Institut Polytechnique de Paris and Inria Saclay

Jan. 2021 - March 2024

Palaiseau (FR)

Computer Science Master degree, Artificial Intelligence track

Université Claude Bernard Lyon I

2018 - 2020

Villeurbanne (FR)

Computer Science Bachelor degree

Université Claude Bernard Lyon I

2015 - 2018

Villeurbanne (FR)

Professional service

Teaching

2022-2023

- Intern co-supervisor: 4 Polytechnique students (2023 and 2022), 1 IMT student (2022)
- Teaching Assistant: Machine Learning (spring 2023), Algorithms (fall 2022), Object-Oriented Programming (spring 2022)

PC and review

2021 - now

- Organization: CMLS 2025 (workshop@ER)
- PC: DBKDA 2025, ADBIS 2025, iSAILS 2025 (workshop@CAiSE), demos@BDA 2025
- Journal: GigaScience 2025, JDIQ 2025, CMC 2025, PLOS ONE 2024, BMC Medical informatics and Decision Making 2024, Scientific Reports 2024
- Conference (external reviewer): ICDE 2025, CIDR 2023, EDBT 2023, ICWE 2021

Talks

2022 - now

- From data to journalism (international high school of Palaiseau - 2024)
- Artificial Intelligence: a tool for journalism (French media development agency - 2023)
- Data integration research: the case of ConnectionLens (young female scientists meeting - 2022)

Other services

2021 - now

- Participate to the “FAIRification of genomic annotations” and “DEI” working groups (2023-now)
- Manage the CEDAR team seminars (2021-2024)
- Co-develop the software to manage Polytechnique students internships (2021)

Skills

Programming languages

Java, Python, Bash, Prolog, C/C++, PHP

Frameworks

Scikit-learn, Flask, Bootstrap, Spring, D3.js,

Web

HTML, CSS, JavaScript

Databases

MySQL, PostgreSQL, MongoDB

Tools

Git, Shell, Maven, JMerise, UMLet

Languages

French (native), English (professional)

Soft skills

dynamism, teamwork, pedagogy, vulgarization

Leisures

I currently enjoy doing puzzles. During my university years, I enjoyed being a team leader and Web developer for the Computer Science challenge “Nuit de l’Info”.