

Nelly Barret

PhD student in CEDAR team at INRIA

75 rue du 1er mars 1943, 69100 Villeurbanne
07 69 90 95 01 — nelly.barret@inria.fr — Driver licence
nellybarret.fr — [nelly-barret](https://inria.fr) — [ORCID](https://orcid.org/)

Education

PhD degree, Computer science track

2020 - 2023

Polytechnique Institute

Palaiseau

- subject: “scalable and expressive integration of heterogeneous datasources”
- technologies: Java, SQL, Postgres and Maven
- my thesis is based on the [ConnectionLens](#) project
- PhD student activities: giving talks about my thesis and the research to various audiences, attending and present my research in conferences, teaching (Java course), scientific and transversal trainings

Computer Science Master, Artificial Intelligence track

2018 - 2020

University of Lyon

Villeurbanne

- courses on cognition and intelligence, machine learning, neural networks and computer environments for human learning
- projects: data mining for predicting cardiovascular diseases, [data visualization](#) about the evolution of programming languages, analysing activity from Twitch users

Computer Science Bachelor

2015 - 2018

University of Lyon

Villeurbanne

- courses on web programming, databases, algorithmic and networks
- project: [interactive map of Lyon](#) for recommending pubs and restaurants with Google and Yelp

Scientific Bacallaureate

2012 - 2015

High school Blaise Pascal

Charbonnières-les-bains

Scientific publications

Toward Generic Abstractions for Data of Any Model

Oct. 2021

BDA - Gestion de Données : Principes, Technologies et Applications

Paris, FR

Potential users need to understand a dataset in order to decide if it is useful for their goal. While some datasets come with a schema and/or documentation, this is not always the case. We present our ongoing work on a novel data summarization approach, which applies on several data formats and computes a description meant for humans, i.e. expressive and compact.

Facilitating Heterogeneous Dataset Understanding

Oct. 2021

BDA - Gestion de Données : Principes, Technologies et Applications

Paris, FR

This PhD student paper presents a system capable to create compact and expressive descriptions out of any dataset, a challenge raised by the big and open data. The two main use-cases of this work is to help users without advanced IT skills discover a new dataset and to help researchers to provide documentation with their datasets.

Predihood: an open-source tool for predicting neighbourhoods' information Sept. 2020
JOSS - Journal of Open Source Software *Journal*

This paper describes the main features of Predihood, a tool for predicting metrics for neighbourhoods (to help people finding a living place or to help scientists to evaluate the pollution) in a user-friendly interface. Moreover, Predihood proposes a generic and easy-to-use programming structure for machine learning algorithms, based on Scikit-learn algorithms. Thus, experts can implement hand-made algorithms and run experiments in it.

Predicting the environment of a neighborhood: a use case for France July 2020
DATA - International Conference on Data Science, Technology and Applications *Online*

This article presents the definition of six environment variables for simply describe the environment of a neighbourhood. Values of these variables are predicted by integrating heterogeneous data sources on which machine learning techniques are applied.

Spatial entity matching with GeoAlign Nov. 2019
ACM SIGSPATIAL - International conference on advances in GIS *Chicago, USA*

Spatial entity matching aims at detecting correspondences between two entities that represent the same point of interest (POI). We propose a customizable formula for detecting and merging these correspondences and an estimation of quality on detected correspondences.

À la recherche du quartier idéal Jan. 2019
EGC - Conférence francophone sur l'Extraction et la Gestion des Connaissances *Metz, FR*

In this article, we present a tool which facilitates the comparison between neighbourhoods. We exploit many indicators (integrated from different data sources) for differentiate neighbourhoods and some algorithms to either recommend neighbourhood in comparison with a departure neighbourhood or group similar neighbourhoods.

Research experience

Full-time contract: classification of complex objects in ConnectionLens Oct. - Dec. 2020
INRIA lab *Palaiseau*

- study and understand the classification process of complex objects stored by Connection Lens graphs
- vulgarise and present work to other team members
- keep informed by attend conferences and seminars

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

- draw up a state of the art about prediction techniques
- propose solutions to scientific challenges: an algorithm for selecting criteria, tacking into account the distribution of criteria to improve prediction
- develop a generic interface for tuning prediction algorithms and for neighbourhood cartographic visualization
- present work to different listeners such as the social science researchers in the project
- write the *Prediction* section of a scientific publication

Master project: matching and fusion of geographic entities with GeoAlign Jan. - June 2019
LIRIS lab *Villeurbanne*

- integrate heterogeneous cartographic data from Geonames, Bing, Here and OSM
- match providers schemas into a unique one

- propose a customizable formula to detect correspondences
- estimate the quality of detected correspondences in an automatic way
- merge automatically detected correspondences
- develop an interface to detect and merge correspondences between point of interest and include into it an estimation of quality

Internship: comparison/recommendation of neighbourhoods with VizLiris May - July 2018
LIRIS lab Villeurbanne

- draw up a state of the art on recommending techniques
- integrate data from heterogeneous data sources
- use prediction algorithms to recommend neighbourhoods
- use clustering algorithms to classify neighbourhoods
- develop an interface which facilitates comparison and recommendation of neighbourhoods in France

Work experience

Seasonal contracts Summers 2016 - 2019

- factory worker at Metaldyne - check quality of BMW pulleys
- temporary worker at University of Lyon - register students, animate the back-to-school amphitheatres
- vendor at Bershka - work in the shop (shop, cabins, stock) and help clients

Volunteering - GAF (female gymnastics) instructor 2010 - 2017
Le Cran de Tassin Tassin

- lead a gymnastics class with 25 young gymnasts in competition
- learn about animation and gymnastic skills (with diplomas)
- participate in the life of the association through the organization of events (gala and performances) and meetings

Skills

Programming languages	Python, Java, Bash, Prolog, C/C++, PHP
Frameworks	Scikit-learn, Flask, Leaflet, Bootstrap, PyTorch, D3.js, Jade, Discord API
Web	HTML, CSS, JavaScript
Databases	MySQL, PostgreSQL, MongoDB
Miscellaneous	L ^A T _E X, Office family
Tools	Git, Shell, Maven, JMerise, UMLet
Languages	French (native), English (professional)
Certifications	Google Digital Active certification, C2I
Other	dynamism, teamwork, written expression

Activities

- participation in the 3 editions of the [Nuit de l'Info](#) as leader of my team and Web developer
- side-car demonstrations on tracks
- formerly, gymnastics in competition