

# URSZULA CZRWINSKA

## DATA SCIENTIST

 Nation, Paris  
 +33698212083  
  urszulaczerwinska  
 urszula.czerwinska@cri-paris.org

### Projects portfolio:

<https://urszulaczerwinska.github.io/works>

French and Polish citizenship  
Driving license type B

## SKILLS

### Technical

Python, SQL, Git,  
tensorflow, pytorch, spacy

OS: Unix/Mac, Windows

### Languages

French, English, Polish  
Spanish, Russian

## PROFESSIONAL EXPERIENCE

### NLP engineer at French Supreme Court

*Senior Data Scientist, june 2020 – currently, Cour de Cassation, Paris*

- LM, NER, Classification (flair, hugging face transformers, pytorch, spacy, allennlp, mongo, elasticsearch)
- Managing data pipelines: DVC
- Data evaluation, models testing & monitoring.

### Machine Learning projects tailored to the needs of our clients from CAC40

*Senior consultant Data Scientist, June 2019 – June 2020, Saegus, Paris*

- **Société Générale:** XAI : user story on algorithms interpretability (SHAP, Alibi, counterfactual predictions, scikitlearn, jupyter notebooks, OOP python code)
- **Servier:** NLP analysis of biotech patents: classification, NER (Elasticsearch, Python: scikitlearn, spacy, flair, tensorflow, pytorch, Azure)
- Show case for tech conférence **EGG PARIS 2019:** keywords extraction from Medium blog posts (Dataiku DSS, Python, spacy, API: flask, js)

### Data management et data strategy for indirect purchasing consultancy

*Lead data scientist, Nov. 2018 – May 2019, JICAP-PERFORMANCE, Paris*

- realized POCs (python) of data extraction from PDFs (text and images)
- formulated a strategy of management et data sharing
- leaded development of the analytics branch of the company

### Data Science applied to biomedical sciences

*Junior researcher, Oct. 2015 –Nov. 2018, Institut Curie, Paris*

- analyzed **large data** (>100Gb) – the biggest referenced study of its kind, using **machine learning, statistical and distributed computing** tools; **improved state of the art** by 5 % creating a new analytical tool (R package); demonstrated **a novel application of an algorithm** (Independent Components Analysis) in oncoimmunology; created **dashboards** (Tableau, RShiny), **reports** (ppt, pdf, HTML) and **data visualizations** (R ggplot2, plotly)
- took part in 6 scientific publications (of whom lead 2) in different teams of 2-15 people
- gave talks (in front of 20 - 1000 people) at 6 **national and international symposiums**, won an award in a professional pitch competition, awards of best oral and poster presentations
- taught **maths, statistics, and programming** at Pharmacology faculty of Paris Descartes (to students in 2<sup>nd</sup> and 5<sup>th</sup> year)

## OTHER

- wrote chapter 'interpretability of ML models' for a book "Data Science in Tourism", Springer 2020
- co-organized Meetup events « Women in Healthcare and Data Science », 2018-2019
- gave a seminar at Créalise, Réunion, data jobs at Epitech, Réunion, mai 2019
- cofounded a biotech start-up, Apr. 2015- Sept. 2015

## EDUCATION

**Ph.D. in bio-mathematics**, Institut Curie, Paris Descartes, Centre de Recherches Interdisciplinaires (CRI);

**complementary courses:** SciencesPo Paris, L'Ecole Polytechnique, Coursera, 2015-2108

**Master**, Interdisciplinary approaches to life sciences, CRI, Paris Diderot, highest honors, 2013- 2015

**Double bachelor**, Biology-Mathematics, UPMC, Paris VI, honors, 2010-2013

**University exchange**, National University of Singapore, Singapore, 2012-2013

**Bilingual French-Polish high school**, Warsaw, Poland, high honors, 2010