



João P C Bertoldo

26, Brazilian

[Online CV](#)

[joaopcbertoldo.github.io](https://github.com/joaopcbertoldo)

April, 2021

(+33) 06 25 51 62 57

joao.bertoldo@mines-paristech.fr

[injoaopcbertoldo](#)  0000-0002-9512-772X

EDUCATION

Paris-Dauphine – PSL University ^(a)	Paris, France
MSc in Artificial Intelligence, Systems, Data (IASD)	2019 – 2020
<i>Subjects: Machine Learning, Deep Learning, Optimization, Image Analysis</i>	
MINES ParisTech- PSL University ^(b)	Paris, France
MS in Executive Engineering <i>Minor: Data Science</i>	2017 – 2020
<i>Subjects: Software Engineering, Databases, Probability, Statistics, Operations Research</i>	
ENSIAME - INSA Hauts-de-France ^(c)	Valenciennes, France
Mechatronics Engineering (1-year academic exchange)	2015 – 2016
<i>1-year scholarship for an international academic exchange programme</i>	
University of São Paulo (USP) ^(d)	São Carlos, Brazil
Mechatronics Engineering (5-year degree)	2013 – 2015, 2016 – 2017

RESEARCH

The European Synchrotron Radiation Facility (ESRF) ^(e)	Grenoble, France
MINES ParisTech- PSL University	Paris, France
Research Engineer	since Mar. 2021
<i>Working on X-ray diffraction tomography (DCT) semantic segmentation.</i>	
MINES ParisTech- PSL University ^(f)	Paris, France
Research Intern	2020 – 2021 (6 mos.)
<i>Worked on semantic segmentation neural networks (2D and 3D U-Net) applied to X-ray tomography.</i>	
<i>Presented results at the 44th edition of the Journée ISS France.</i>	
<i>Lectured about U-Nets in a week-course (CVML 2021) at the PSL University.</i>	
Mindsay ^(g)	Paris, France
Research Intern	2020 (5 mos.)
<i>Evaluated explainability methods (LIME, SHAP, Anchors) applied to layperson-maintained Natural Language Processing (NLP) classification models.</i>	

PROFESSIONAL EXPERIENCE

Datadog ^(h)	New York, USA
Data Scientist Intern	2019 (7 mos.)
<i>Developed algorithms for time series and tag analysis, and contributed to a knowledge graph project.</i>	
<i>Maintained high-impact production Python codebase and optimized an automated test pipeline.</i>	
Kelda Drilling Controls ⁽ⁱ⁾	Porsgrunn, Norway
Software Engineer Intern	2018 – 2019 (8 mos.)
<i>Designed a web application for time series supervised annotation.</i>	
Stone Co. ^(j)	Rio de Janeiro, Brazil
Backend Developer / Data Scientist Intern	2017 – 2018 (3 mos. / 2 mos.)

*Developed a .NET web Rest API for a MongoDB and lectured about Design Patterns.
Explored transactional data to predict churn risk with decision trees-based methods.*

University of São Paulo (USP) ^(k)

Teaching assistant (part-time)

São Carlos, Brazil

2015 (6 mos.)

Elaborated and conducted workshops on numerical analysis using MATLAB.

PROJECTS

tomo2seg (f,e)

since Sep. 2020

Developed an X-ray tomography (XCT) semantic segmentation tool using U-Nets.

The [data](#) and the [models](#) are publicly available on Zenodo. A [tutorial](#) is available on GitHub(WIP).

Applied it during an experiment at the Soleil Synchrotron, France.

→ [Link to tomo2seg on GitHub](#)

pymdr (a)

2020 (1 mo.)

Python implementation of *Mining Data Records*

([liuMiningDataRecords2003](#), [liuMiningDataRecords2003](#))

→ [Link to pymdr on GitHub](#)

Correlations (h)

2019 (7 mos.)

Contributed to build *Correlations*: an automated debug tool for complex infrastructures capable of finding correlated accidents in time series.

→ [Link to Correlations on Datadog's website](#)

Bivrost (i)

2018 (8 mos.)

Maintained and implemented new features in Bivrost: a web application for high-resolution time series visualization used for real-time drilling systems monitoring.

→ [Link to Bivrost's page on Kelda's website](#)

Sunlight simulator (b)

2017 – 2018 (9 mos.)

Lead a project for 6 mos. to build a robotic cable-controlled sunlight physical emulator.

Developed a 3D motion digital twin in Python.

→ [Link to minlight on GitHub](#)

Smart coffee machine (d)

2016 (4 mos.)

Designed a coffee machine controlled over the internet.

Developed a web server in .NET (C#) and an Arduino controller (C++).

→ [Link to mkafeina on GitHub](#)

PUBLICATIONS

SKILLS

Languages

English Fluent

Obtained TOEFL 117/120 (2020) and worked in English-speaking companies for more than 1 year.

French Fluent

Read subjects in French for 3 years and worked in French-organizations for 1 year.

Portuguese Native

Turkish Beginner

Programming

Python Fluent

Data processing/viz. Numpy, Scipy, Pandas, IPython, Matplotlib, Seaborn.

Machine/Deep learning Keras, TensorFlow, Scikit-learn.

Prototyping Streamlit, Jupyter Lab/Notebook.

Miscellaneous Anaconda, Virtual Environments, nose, unittests.

Other programming languages

MATLAB, JavaScript (React, Redux), C/C++ (basics), Java (basics).

Databases

PostgreSQL, MongoDB, InfluxDB.

Computer skills (others)

IDEs Pycharm, Visual Studio Code (VSCode).

Miscellaneous Git, GitHub/GitLab, Linux, Bash/Zsh, SSH, Markdown, Weka.

Tomography ImageJ (Fiji), Avizo.

Content creation/management

LaTeX Fluent. Wrote [Master's internship report](#) in LaTeX.

Google Suite (Docs, Slides, Sheets) Fluent. Wrote [Master's defense](#) presentation with Google Slides.

Zotero Extensively using for more than 1 year.

INTERESTS

TALKS

TEACHING
