

PEDRO OLIVEIRA cpdomina@gmail.com | <http://cpdomina.net>

Software Engineer specialized in data analysis. Interested in machine learning, natural language processing, knowledge graphs, recommendation/discovery/search engines, and computational creativity.

WORK EXPERIENCE

Since Feb 2021 AI Consultant, *Self-employed* (remote)

Worked with several startups on AI-related projects, with an emphasis on Large Language Models and advanced Natural Language Processing.

Nov 2018–Dec 2020 Researcher/Musician, *Mountain Pass*

Researched and developed a set of artificial intelligence tools for automating elements of the music production process. Produced and self-released an electronic music album.

Dec 2016–Oct 2018 Senior Software Engineer, *Stardog* (remote)

Designed and developed the machine learning and natural language processing components of Stardog, a knowledge graph database. Also responsible for maintenance and performance improvements on the search, spatial, and document management components, plus frequent support to the sales team.

Jun 2016–Dec 2016 Data Scientist, *wonderif*, Porto, Portugal

Founder of a one-man data analysis company. Helped companies explore and understand their data for content marketing purposes.

Feb 2013–July 2014 Data Engineer, *Disqus*, San Francisco, California

Large-scale data analysis and engineering: built algorithms and data pipeline for a computational advertising system; improved and maintained the content recommendation engine; built an internal search and discovery system.

Dec 2011–Feb 2013 Research Engineer, *Technicolor Research*, Palo Alto, California

R&D on recommendation engines. Mostly focused on using external signals, such as mood and context, to improve movie recommendation and discovery.

Feb 2011–Nov 2011 Founder, *Oriente*, Portugal

Founder of Oriente, a social media analytics platform for the Portuguese-speaking market. Dealt with all the components of the business, from research and development to marketing and customer acquisition.

Nov 2010–Feb 2011 Software Engineer, *Clark&Parsia* (remote)

Developed and improved several key parts (indices and runtime statistics) of their database, Stardog.

Nov 2009–Nov 2010 Artificial Intelligence R&D Intern, *Clark&Parsia*, Washington, DC

Added advanced features to existing products using machine learning, natural language processing, social network analysis, and statistical inference techniques. Helped design and build the first prototype of Stardog, a knowledge graph database.

Jul 2006–Oct 2009 Associate Researcher, *Center for Informatics and Systems of the University of Coimbra*, Portugal

Participated in several research projects in the fields of neural networks, multi-agent systems, natural language user interfaces, social network analysis, semantic web, and statistical relational learning.

EDUCATION

2004-2009 BSc+MSc in Informatics Engineering, University of Coimbra, Portugal

PUBLICATIONS

S. Salamatian, A. Zhang, F. Calmon, S. Bhamidipati, N. Fawaz, B. Kveton, Pedro Oliveira, and N. Taft: "Managing your private and public data: Bringing down inference attacks against your privacy", *IEEE Journal of Selected Topics in Signal Processing*, 9(7):1240-1255, 2015 [[pdf](#)]

S. Salamatian, A. Zhang, F. Calmon, S. Bhamidipati, N. Fawaz, B. Kveton, Pedro Oliveira, and N. Taft: "How to hide the elephant – or the donkey – in the room: Practical privacy against statistical inference for large data", *IEEE Global Conference on Signal and Information Processing*, Austin, Texas, December 2013 [[pdf](#)]

Pedro Oliveira, P. Gomes: "Learning and Reasoning about Uncertainty in the Semantic Web", *14th Portuguese Conference on Artificial Intelligence*, Aveiro, Portugal, October 2009 [[pdf](#)]

Pedro Oliveira, "Probabilistic Reasoning in the Semantic Web using Markov Logic", MSc Thesis, Department of Informatics Engineering, University of Coimbra, Portugal, July 2009 [[pdf](#)]

Pedro Oliveira, P. Gomes: "Instance-based Probabilistic Reasoning in the Semantic Web", *18th International World Wide Web Conference*, Madrid, Spain, April 2009 [[pdf](#)]

R. P. Costa, Pedro Oliveira, G. Rodrigues, B. Leitão and A. Dourado: "Epileptic Seizure Classification Using Neural Networks with 14 Features", *12th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems*, Zagreb, Croatia, September 2008 [[pdf](#)]

PATENTS

E. Sirin, M. H. Grove, K. G. Clark, J. M. Balint, Pedro Oliveira, P. Klinov: "Multi-Source-Type Interoperability and/or Information Retrieval Optimization", [WO2019241293](#), December 2019

E. Sirin, M. H. Grove, K. G. Clark, Pedro Oliveira : "System and method for providing prediction-model-based generation of a graph data model", [US20190384863](#), June 2018

N. Fawaz, S. Salamatian, A. Zhang, F. Calmon, S. Bhamidipati, Pedro Oliveira, N. Taft, and B. Kveton: "Privacy against inference attacks under mismatched prior", [US20160006700](#), January 2016

J. Bolot, Pedro Oliveira: "Method and system for targeted crowdsourcing", [US20150025963](#), January 2015

N. Fawaz, S. Salamatian, A. Zhang, F. Calmon, S. Bhamidipati, Pedro Oliveira, N. Taft, and B. Kveton: "Privacy against interference attack for large data", [WO2014123893](#), August 2014

Pedro Oliveira: "Context-based Content Recommendations", [WO2014051644](#), April 2014

SELECTED CONTENT

[A Primer on Generative Machine Learning Models for Music Production](#), 2022

[How I Made an Album Without Knowing Anything About Music](#), 2020

[Entity Linking in the Knowledge Graph](#), 2018

[Learning to Predict](#), 2017

[Are news outlets selectively negative about certain subjects?](#), 2016

[Is there more sex, violence, and profanity in today's movies?](#), 2016

[Hillary vs Trump: Who is the most emotional during public speeches?](#), 2016

MORE INFORMATION

[Personal Website](#)

[Linkedin](#)

[Github](#)