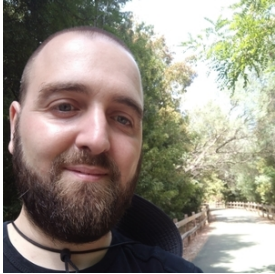# josédevezas

software engineering and data science

contact

Porto, Portugal

joseluisdevezas@gmail.com
(preferred method of contact)

http://www.josedevezas.com
linkedin://jldevezas

languages

portuguese (native)
english (fluent)

programming

♥ Scala Python R SQL
Java C# Rust Gremlin
HTML CSS JavaScript

education

2016–Now	**Doctoral Program** in Informatics	U.Porto, UA, UMinho

Graph-Based Entity-Oriented Search

Doing a thesis in information retrieval, while using knowledge from areas of specialization such as information extraction, network science or machine learning.

2003–2010	**Master** in Informatics and Computing Engineering	U.Porto

Classification: 14 (out of 20)

Link Ecosystem of the Portuguese Blogosphere

This thesis explored the blogosphere as an ecosystem of blogs comprised not only of a set of posts and their individual content, but also of the interactions established through the hyperlinks connecting the posts. Based on this study, we discovered a correlation between the blog's centrality and the length and number of posts.

Thesis Classification: 18 (out of 20)

2000–2003	**Technological Specialization** in Informatics	Colégio de Gaia

Classification: 18 (out of 20)

experience

2017–Now	**FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO**	Porto, Portugal

Invited Assistant Professor

Teaching one semester about social network analysis and another one about databases.

2016–Now	**INESC TEC**	Porto, Portugal

Researcher / MAP-i PhD Student

MAP-i PhD student and FCT grant holder, working on entity-oriented search. Former FourEyes researcher with a focus on studying information consumption on social media. Provided support to the research group through data engineering, by developing the infrastructure for Twitter data collection and analysis, as well as news collection and archiving.

2015–Now	**LABORATÓRIO DE SISTEMAS DE INFORMAÇÃO/U.PORTO**	Porto, Portugal

Researcher / MAP-i PhD Student

MAP-i PhD student and FCT grant holder, working on entity-oriented search. Research and development of the ANT entity-oriented search engine. Particularly focused on query analysis methodologies, as well as the construction of specialized indexes and ontologies for efficient and effective question-answering through contextual widgets. The produced system was also aimed at providing a platform able to support the academic community with the experimentation of search engines backed by linked data.

| 2014–2015 | **INTERRELATE** | Porto, Portugal |

Software Engineer / Data Scientist

Data Science:

- Collection and analysis of several types of data, mainly originating from the web, particularly from social media.
- Highly proficient in R, Python and SQL (from data collection to reporting), including topic tracking, outlier and anomaly detection, statistical smoothing, regression, classification, clustering and network analysis.
- Skilled in dynamic reporting.
- Developed several data visualization widgets.
- Analyzed social networks and other graphs.

Software Engineering:

- Software design of web applications and analytics pipelines.
- Technology evaluation and selection.
- Supporting the team with new technologies or languages.
- INTERRELATE Backoffice:
  - Devising strategies for automatic and periodical data collection and information extraction from unstructured user-generated content;
  - Systems integration, including team coordination (4 person project).

Programming:

- Worked in nearly every company project, using various different languages.
- Developed several web applications and web services.
- INTERRELATE Backoffice:
  - Web crawler development;
  - Development of data visualization widgets for internal statistics;
  - Front-end design and development.

Sysadmin:

- Manage the company servers, including security, backups and product deploys.
- Maintain firewall rules and SMTP server.
- DNS configurations (including SPF, for SPAM control).
- Nginx configurations.
- Company services maintenance (internal and external).
- Log monitoring and bug reporting.
- Real-time e-mail log error monitoring for critical systems and web service mobile monitoring for downtime or connection issues.
- Support Ticket System configuration.
- PayPal payment services.

| 2012–2013 | **LABORATÓRIO SAPO/U.PORTO** | Porto, Portugal |

Researcher / Developer

Research and development of music discovery and recommendation techniques supported on a graph-based neighborhood approach, as well as on latent factor models based on matrix factorization algorithms. Development of methodologies to combine knowledge from different dimensions, including not only user profile, content and context features, but also social features, such as community structure. Related tasks required work on data mining, multimodal network analysis and community detection, and information retrieval.

| 2011–2012 | **CRACS/INESC TEC** | Porto, Portugal |

Researcher / Developer

Enhancement, adaptation and implementation of state of the art algorithms for network analysis, data mining and data visualization, in the context of intelligent information systems for cyber journalism, as part of the Breadcrumbs Project (UTA-Est/MAI/0007/2009).

| 2010-2011 | **LABORATÓRIO SAPO/U.PORTO** | Porto, Portugal |

Researcher / Developer

Work in the areas of data mining, network analysis, community detection, information retrieval, real-time data analytics and data visualization.

| 2002-2003 | **NÚCLEO PORTUGUÊS DE FREEBSD** | Coimbra, Portugal |

Documentation / Programming

Translation of the official FreeBSD documentation to portuguese, creation of how-to guides in portuguese, porting open source applications to FreeBSD and programming and content management of the hint subsystem.

awards

| 2012 | **Certificate of Merit** The 2012 IAENG International Conference on Data Mining and Applications |

Awarded to the best papers published in the conference.

communication skills

| 2017 | **Presentation** | Symposium on Languages, Applications and Technologies |

Presented a complete pipeline, from data acquisition, passing through information extraction and the automatic construction of knowledge base, to an information retrieval implementation.

| 2013 | **Presentation** | DEI–FEUP |

Did a short presentation on the Juggle project for the department, in the context of DEI Talks.

| 2013 | **Presentation** | SAPO |

Presented the Juggle project, including the developed graph-based recommendation algorithms, for the SAPO technical team.

| 2012 | **Presentation** | International Conference on Knowledge Discovery and Information Retrieval |

Presented an interactive news clips visualization tool for applications in journalistic research and knowledge discovery.

| 2012 | **Presentation** | IAENG International Conference on Data Mining and Applications |

Presented a methodology, based on community detection in multidimensional networks, for the creation of news context from a folksonomy of web clipping.

| 2011 | **Poster** | International AAAI Conference on Weblogs and Social Media |

Presented research work on using the h-index to estimate blog authority.

| 2010 | **Presentation** | Laboratório SAPO/U.Porto |

Presented Ciclope project for the Portugal Telecom's CEO and the media, as part of the inaugural presentation of Laboratório SAPO/U.Porto.

| 2010 | **Poster** | KDD Workshop on Social Media Analytics |

Presented the blog popularity study conducted during my Masters degree.

interests

professional: search engines, data analysis, web development, algorithms and data structures for low level information manipulation, visualization **personal:** music, television, guitar, cooking

additional information: Please check out my personal website, where I've shared some of my research output: http://josedevezas.com/#/research#publications

collaborations

2017/2018   **Member of the program committee**
Complex Networks 2017

               http://complexnetworks.org

2016/2017   **Master's thesis co-supervision**
Named entity extraction from Portuguese web text

               http://hdl.handle.net/10216/106094

2015/2016   **Master's thesis co-supervision**
Exploring the Sea: Heterogeneous Geo-Referenced Data Repository

               http://hdl.handle.net/10216/85612

publications

journal articles

Graph-Based Entity-Oriented Search: Imitating the Human Process of Seeking and Cross Referencing Information
José Devezas and Sérgio Nunes
ERCIM News, Special Theme: Research and Innovation *111.3 (2017)*. 2017

The Community Structure of a Multidimensional Network of News Clips
José Luís Devezas and Álvaro Reis Figueira
International Journal of Web Based Communities *9.3 (2013) pp. 411–429. 2013*

Finding Language-Independent Contextual Supernodes on Coreference Networks
José Devezas and Álvaro Figueira
IAENG International Journal of Computer Science *39.2 (2012) pp. 200–207. 2012*

book chapters

Álvaro Figueira, José Devezas, Nuno Cravino, and Luis-Francisco Revilla. "Creating and Analysing a Social Network Built From Clips of Online News". In: *Information Systems and Technology for Organizations in a Networked Society*. IGI Global, 2012.

conference articles

Information Extraction for Event Ranking
José Devezas and Sérgio Nunes
Proceedings of the 6th Symposium on Languages, Applications and Technologies (SLATE 2017), 2017, Vila do Conde, Portugal

Benchmarking Named Entity Recognition Tools for Portuguese
André Pires, José Devezas, and Sérgio Nunes
Proceedings of Simpósio de Informática (INForum 2017), 2017, Aveiro, Portugal

Index-Based Semantic Tagging for Efficient Query Interpretation

José Devezas and Sérgio Nunes

Proceedings of the 6th International Conference of the CLEF Initiative (CLEF 2016), 2016, Évora, Portugal

Exploring a Large News Collection Using Visualization Tools

Tiago Devezas, José Devezas, and Sérgio Nunes

Proceedings of the First International Workshop on Recent Trends in News Information Retrieval co-located with 38th European Conference on Information Retrieval (ECIR 2016), 2016, Padua, Italy

Large-scale Crossmedia Retrieval for Playlist Generation and Song Discovery

Filipe Coelho, José Devezas, and Cristina Ribeiro

Proceedings of the 10th International Conference in the RIAO series (OAIR 2013), 2013

Using the Overlapping Community Structure of a Network of Tags to Improve Text Clustering

Nuno Cravino, José Devezas, and Álvaro Figueira

Proceedings of the 23rd ACM Conference on Hypertext and Social Media (HT 2012), 2012, Milwaukee, WI, USA

Creating News Context From a Folksonomy of Web Clipping

José Devezas, Henrique Alves, and Álvaro Figueira

Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2012 (IMECS 2012), 2012, Hong Kong

Studying a Personality Coreference Network in a News Stories Photo Collection

José Devezas, Filipe Coelho, Sérgio Nunes, and Cristina Ribeiro

Lecture Notes in Computer Science: Proceedings of the 34th European Conference on Information Retrieval (ECIR 2012), 2012, Barcelona, Spain

Interactive Visualization of a News Clips Network: A Journalistic Research and Knowledge Discovery Tool

José Devezas and Álvaro Figueira

Proceedings of the 4th International Conference on Knowledge Discovery and Information Retrieval (KDIR 2012), 2011, Barcelona, Spain

Using the H-index to Estimate Blog Authority

José Devezas, Sérgio Nunes, and Cristina Ribeiro

Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011), 2011, Barcelona, Spain

FEUP at TREC 2010 Blog Track: Using h-index for blog ranking

José Devezas, Sérgio Nunes, and Cristina Ribeiro

The Nineteenth Text REtrieval Conference Proceedings (TREC 2010), 2010, Gaithersburg, MD, USA

Studying Blog Features over Link Popularity

José Devezas, Cristina Ribeiro, and Sérgio Nunes

Proceedings of the First Workshop on Social Media Analytics (SOMA 2010), held in conjunction with the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2010), 2010, Washington, DC, USA

thesis reports

José Devezas. "Link Ecosystem of the Portuguese Blogosphere". Master's Thesis. Faculdade de Engenharia da Universidade do Porto, 2010.

research reports

Entity-Oriented Search - FEUP InfoLab Report

José Devezas

Tech. rep., 2016

An Overview of the Graph Database Paradigm - Breadcrumbs Project

José Devezas

Tech. rep., 2011

Overlapping Community Detection - Laboratório SAPO/U.Porto Report

José Devezas, Sérgio Nunes, and Cristina Ribeiro

Tech. rep., 2011