

# José Devezas, Ph.D.

data → information → knowledge → search & analytics



[joseluisdevezas@gmail.com](mailto:joseluisdevezas@gmail.com)

<https://josedezas.com>

LinkedIn: [jdezaz](#)

## Location

Porto, Portugal

## Languages

Portuguese (native)

English (fluent)

## Main Skills

Graph Analytics

Neo4j

Gremlin

Lucene

SQL

Python

R

Scala

Java

Vue.js

## Areas

Network Science

Search Engineering

Knowledge Engineering

Data Science

Data Engineering

Software Engineering

## Education

2016–2021 **Ph.D.** in Computer Science (MAP-i)

U.Porto, UA, UMinho

Curricular grade: 19 (out of 20)

Thesis: [Graph-Based Entity-Oriented Search](#)

Two main systems were developed:

- **ANT**, an entity-oriented search engine for the University of Porto.
- **Army ANT**, a workbench for innovation in entity-oriented search.

**Publications:** <http://josedezas.com/academy/publications/>

**Projects:** <http://josedezas.com/academy/projects/>

**Slides:** <http://josedezas.com/academy/slides/>

**Languages:** Python, R, Java, Gremlin; **Frameworks:** Flask, Vue.js; **Storage:** Lucene, Neo4j, PostgreSQL

2003–2010 **M.Sc.** in Informatics and Computing Engineering (MIEIC)

U.Porto

Final grade: 14 (out of 20), Thesis grade: 18 (out of 20)

Thesis: [Link Ecosystem of the Portuguese Blogosphere](#)

**Languages:** Perl, R; **Libraries:** igraph; **Storage:** BerkeleyDB, Lucene

2000–2003 **Technological Specialization** in Informatics

Colégio de Gaia

Classification: 18 (out of 20)

## Experience

2023–2024 **Pasabi**

Portugal / United Kingdom (Remote)

*Senior Data Scientist*

Worked on fake review detection, particularly scam detection. Designed an R&D lifecycle workflow that established the methodology for our data team. Day-to-day tasks revolved around text classification, graph-based algorithms and classification, data analysis and visualization, prompt engineering, data acquisition approaches, data lakehouse architecture and tooling, mostly running on top of cloud infrastructure.

**Languages:** Python, Typescript, SQL, Terraform/HCL; **Frameworks:** FastAPI, Kedro, MLflow; **Storage:** Neo4j, PostgreSQL, MongoDB, Iceberg; **Software:** Jupyter, Elasticsearch/Kibana; **Services:** GCP, AWS, ChatGPT; **Skills:** R&D Processes, Text Classification, Graph Data Science, ELT, Prompt Engineering

2021–2023 **HEX DATUM OÜ**  
*Founder CEO*

Tallinn, Estonia (Remote)

Hex Datum OÜ offered software-as-a-service (SaaS), as well as consulting services in the area of intelligent information systems. Our flagship product was a crypto analytics web application called Indicator Dash (indicatordash.app).

**Languages:** Python; **Frameworks:** Flask, Vue.js, Tachyons; **Storage:** PostgreSQL, JanusGraph, Redis; **Software:** Jupyter; **Skills:** ETL, Financial Indicators, Data Visualization, Time Series Forecasting, Transformers, LSTMs, Indexing and Search, Sentiment Analysis

2015–2021 **CSIG, INESC TEC & INFOLAB, UNIVERSITY OF PORTO**  
*Researcher / PhD Student*

Porto, Portugal

- MAP-i PhD student and FCT grant holder, working on entity-oriented search.
- FourEyes project researcher with a focus on studying information consumption on social media.
- Data engineer responsible for developing infrastructure for Twitter data collection and analysis, and news collection and archival.
- Research and development of the **ANT** entity-oriented search engine.
  - Construction of specialized indexes and ontologies for efficient and effective question-answering through contextual widgets.
- Development of the open source project, **Army ANT**, as a platform for research and experimentation of search engines backed by linked data.

**Languages:** Java, Gremlin, R, Python; **Frameworks:** Scrapy, SQLAlchemy, Flask, Vue.js, Spectre.css, Bulma; **Libraries:** igraph, NetworkX, Pandas; **APIs:** Twitter; **Storage:** PostgreSQL, MongoDB, Lucene, Neo4j; **Software:** Jupyter; **Skills:** Scientific Writing, Experimental Design, Indexing and Search, Web Crawling and Scraping, Social Media Analytics

2017–2018 **FACULTY OF ENGINEERING, UNIVERSITY OF PORTO**  
*Invited Assistant*

Porto, Portugal

Lectured two semesters on:

- **Social network analysis**, for the Information Management in Social Networks course, of the Master in Information Science.
- Relational and document databases, for the Databases course, of the Master in Informatics and Computing Engineering.

**Storage:** SQLite, MongoDB; **Software:** Gephi, KNIME; **Skills:** Educational Content Creation, Slide Presentation, Student Assessment

2014–2015 **INTERRELATE**  
*Software Engineer / Data Scientist*

Porto, Portugal

INTERRELATE was a data analytics and social media monitoring and management company. It provided services based on data retrieval, entity detection, sentiment analysis, data visualization, text mining and natural language processing, event and topic detection, and influencer detection.

**Languages:** Python, R; **APIs:** Twitter, Facebook, PayPal; **Services:** SMTP, DNS, Nginx, Support Tickets; **Storage:** PostgreSQL; **Skills:** Social Media Analytics, Web Crawling and Scraping, Anomaly Detection, Statistical Smoothing, Regression, Classification, Clustering, Graph Analytics, Automated Reporting, Data Visualization, Systems Integration, Web Development, Systems Administration, Backup Strategies, Firewall Rules, Log Monitoring

2012–2013 **SAPO LABS, UNIVERSITY OF PORTO**  
*Researcher / Developer*

Porto, Portugal

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.

**Languages:** Python, Gremlin; **Libraries:** SciPy, PyMF; **Storage:** Neo4j, HDF5, Redis; **Skills:** Recommender Systems, Music Recommendation, Collaborative Filtering, Content Filtering, Graph Analytics

2011–2012 **CRACS, INESC TEC**  
*Researcher / Developer*

Porto, Portugal

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 **SAPO LABS, UNIVERSITY OF PORTO**  
*Researcher / Developer*

Porto, Portugal

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