

# André Mourão, Ph.D.

Lisbon, Portugal  
andre.b.mourao@gmail.com

Software Engineer

github.com/amourao  
linkedin.com/in/andre-mourao

R&D with real world experience in implementing complex Search, Computer Vision and Machine Learning systems.

I've scaled systems from 20 million to 1.8 billion documents, implemented novel pipelines for semantic search, NLP and Computer Vision, and worked with data ranging from smartwatch step data, web-pages, faces, news articles, medical articles, live video, Jira tickets and many others.

## SKILLS

<b>Information Retrieval/Search</b>	Java, Python, Golang, Elasticsearch, Solr, Lucene, Hadoop, trec-eval
<b>Computer Vision</b>	C++, OpenCV, LibAV, LibSVM, FLANN
<b>Data Science and NLP</b>	Python, NumPy, SciPy, scikit-learn, SpaCy, Pika
<b>Other</b>	TypeScript, git, RabbitMQ

## TECHNICAL EXPERIENCE

**Software Engineer - Search** Nov 2023 — Present  
*Weaviate*

- Improving BM25 search features and performance to improve hybrid search

**Senior Software Engineer - Search** Sep 2022 — Jul 2023  
*Mem Labs*

- Combined semantic and keyword search systems into a hybrid system achieving Recall@10 of over 90% on test collection
- Changed data model for low-effort addition of new document types (e.g. email, chat messages, calendar) for search

**Principal Search Engineer** Jul 2021 — Jul 2022  
*Searchable.ai*

- Built hybrid semantic search system from scratch, increasing retrieval performance by 20%
- Used semantic search used as basis for improved Q&A performance and enable document clustering and recommendation

**R&D Engineer** Jan 2020 — Jun 2021  
*Portuguese Web Archive, arquivo.pt*

- Increased Solr image metadata search index 90-fold, from 20 million to 1.8 billion docs
- Responsible for page, image search and user-facing APIs

**Freelancer** Jan 2019 — Jun 2021

- moetsi**: designed a flexible low latency Distributed Video Processing system for Computer Vision at scale

**Post-doc researcher at the COGNITUS Horizon 2020 project** Jan 2018 — Aug 2018  
*Nova SST*

- Created the event plotting module, enabling journalists to find the best photos and videos for news
- Leveraged on social media trends using NLP and unsupervised ML

**Visiting Researcher** Apr 2015 — May 2015  
*University of Texas, Austin*

- Extraction and analysis of smartwatch activity and health data

## EDUCATION

**Ph.D. in Computer Science** 2012 — 2018  
*Supervisor: Prof. João Magalhães* NOVA SST

- Developed an architecture to improve the partitioning of very large scale (over 1 billion) image search for distributed systems.
- Benchmarked on Microsoft Azure cloud, with a difference in load of <10% between nodes.

**M.Sc. degree in Computer Science** 2010 — 2012  
*Supervisor: Prof. João Magalhães* NOVA SST

- Created a SVM model and image features for real-time facial expression recognition
- The model was successfully applied on a competitive game based solely on player's faces.

**B.Sc. degree in Computer Science** 2007 — 2012

## AWARDS AND DISTINCTIONS

**Arquivo.pt Award** – 2<sup>nd</sup> place for the [Revisionista.pt](#) project: Find out how news articles are edited after publication across Portuguese newsrooms 2019

**Text REtrieval Conference (TREC)** – Best *NDCG* TREC FedWeb 2013, Top 5 *P@10* TREC Clinical Decision Support 2014.

**Microsoft Azure Research Award** – Awarded \$40,000 in Microsoft Azure credit to apply on improving index partitioning for distributed search of biomedical multimodal data. 2015 — 2017

**Fraunhofer Portugal Challenge Awarding Research of Practical Utility** – 2<sup>nd</sup> place: M.Sc. category 2013