

# Francisco B. Valero (Fran Valero)

<https://franbvalero.github.io> • [franbvalero@gmail.com](mailto:franbvalero@gmail.com) • (+34) 678 81 24 74  
Barcelona, Spain

## Work Experience

- **Applied Scientist** 03/2022 – Ongoing  
**Amazon.com, Inc.**  
Barcelona, Spain  
Applying NLP and ML techniques to customer reviews in Community Shopping managed by Hani Al-Shater.
- **Data Scientist** 06/2021 – 03/2022  
**Amazon.com, Inc.**  
Barcelona, Spain  
Working on two science projects in the Search Science and AI department managed by Lluís Màrquez:
  - Building an open source search evaluation dataset.  
This work was presented as the dataset for the Amazon Product Search competition in the Conference on *Knowledge Discovery and Data Mining* (KDD 2022).
  - Applying knowledge distillation based on few shot learning for short text generation.
- **Research & Development Intern** 11/2020 – 05/2021  
**Deezer S.A.**  
Paris, France  
Used NLP and ML techniques to encode the podcast's textual metadata for searching supervised by Elena V. Epure.  
The work performed during this internship was accepted as a full paper in the *European Conference on Information Retrieval* (ECIR 2022).
- **Research Engineer Intern** 06/2020 – 09/2020  
**Kustomer, Inc.** (Acquired by Meta Platforms, Inc.)  
Madrid, Spain  
Worked on fine-tune with pre-trained Transformers models for intent detection managed by Víctor Peinado.
- **Research Engineer** 04/2019 – 09/2019  
**Department of Software and Computing Systems, University of Alicante**  
Alicante, Spain  
Used ML models for estimating the quality of the machine translation systems supervised by Mikel L. Forcada for the *EFFORTUNE Spanish project* (TIN2015-69632-R).
- **Research Intern** 01/2018 – 04/2019  
**Department of Software and Computing Systems, University of Alicante**  
Alicante, Spain  
Worked on two research projects supervised by Juan Antonio Pérez-Ortiz and Felipe Sánchez-Martínez:
  - Used neural models as RNN and Seq2Seq for sequence labeling problems.
  - Used cross-lingual word embeddings for word translation.

## Educational Background

- **Master's Degree in Artificial Intelligence, Pattern Recognition and Digital Imaging** 09/2019 – 09/2020  
**Polytechnic University of Valencia**  
Valencia, Spain  
Master's Thesis: *Data augmentation for text classification*  
Advisor: Francisco Casacuberta
- **Bachelor's Degree in Compute Science Engineering** 09/2014 – 02/2019  
**University of Alicante**  
Alicante, Spain  
High Academic Performance Group  
Bachelor's Thesis: *Neural Part-of-Speech tagging*  
Advisors: Juan Antonio Pérez-Ortiz and Felipe Sánchez-Martínez

## Conference papers

- **Francisco B. Valero**, Marion Baranes, Elena V. Epure. Topic Modeling on Podcast Short-Text Metadata. In 44th *European Conference on Information Retrieval (ECIR 2022)*, Stavanger, Norway.
- Chandan K. Reddy, Lluís Marquez, **Fran Valero**, Nikhil Rao, Hugo Zaragoza, Sambaran Bandyopadhyay, Arnab Biswas, Anlu Xing, Karthik Subbian. Shopping Queries Dataset: A Large-Scale ESCI Benchmark for Improving Product Search . In 28th *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2022)*, Washington DC, U.S.

## Grants

- Research Initiation Grant, *Ministerio de Educación Cultura y Deporte de España (MECD)* 2018/2019

## Courses and training

- Deep Learning Specialization, online at Coursera, 2018.

## Languages and Technologies

**Programming Languages:** Python, L<sup>A</sup>T<sub>E</sub>X, Java, C/C++, bash

**Technologies:** NumPy, PyTorch, scikit-learn, NLTK, Transformers, SciPy, pandas, linux terminal, Git, Docker

**Natural Languages:** Native in Spanish and professional working proficiency in English

## Honors and Awards

- 3rd place in PAN@CLEF 2020 Author Profiling Competition

## Reference list

- **Hani Al-Shater** (haniahma@amazon.com)  
Amazon.com, Inc.  
Berlin, Germany
- **Lluís Màrquez** (lluismv@amazon.com)  
Amazon.com, Inc.  
Barcelona, Spain
- **Elena V. Epure** (epure@deezer.com)  
Deezer S.A.  
Paris, France
- **Víctor Peinado** (victor.peinado@kustomer.com)  
Kustomer, Inc.  
Madrid, Spain
- **Mikel L. Forcada** (mlf@dlsi.ua.es)  
University of Alicante  
Department of Software and Computing Systems  
Alicante, Spain
- **Juan Antonio Pérez-Ortiz** (japerez@dlsi.ua.es)  
University of Alicante  
Department of Software and Computing Systems  
Alicante, Spain