



Fábio Lobato

Computer Engineer, Data Scientist

Summary

Highly creative Data Scientist with 10+ years of R,D&I experience, specializing in Artificial Intelligence and Natural Language Processing (NLP) for data-driven decision-making. CNPq Technological Development Fellow with proven experience in ethically-driven dataset creation and annotation for NLP and bioinformatics, specifically within the healthcare and legal domains. I currently work with the Ministry of Science and Technology of Brazil (MCTI) on R,D&I projects evaluation and previously served as an AI Specialist advisor on the G20 Digital Economy Working Group (DEWG), leading contributions to the AI products under the Brazilian presidency.



fabio.lobato@ufopa.edu.br

+55 93 99131 0666

Santarém, Pará, Brazil

[F. Lobato](#)

[F. Lobato](#)

[F. Lobato](#)

[fabiolobato](#)

Portuguese	Native
English	Fluent
Spanish	Advanced

Languages	Python, C/C++, Java
-----------	------------------------

Frameworks	TensorFlow, Keras, NTLK, Pan- das, Numpy, Sklearn, SpaCy, Gensim
------------	--

Language Models	Llama, GPT, BERT, Word2Vec, FastText
--------------------	---

Tools	Hungging Face Datasets, Web Scrap- ing, Doccano
-------	---

Work experience

Senior Posdoctoral Researcher - CNPq funding

Dec 2024 - present

University of São Paulo and Universidad Del País Vasco

- Conducting research on Large Language Models at the Laboratory of Computational Intelligence - LABIC. Collaborating with the *Universidad Del País Vasco* for a three-month period at the Intelligent Systems Group

Lecturer

Jun 2014 - present

Federal University of Western Pará

- Research Group Leader & Laboratory Principal Investigator - Laboratory of Applied Computing

Visiting Researcher

Sep-Dec 2018/2025

Universidad Del País Vasco

- Worked together with the Intelligent Systems Group in two periods of three months in 2018 and 2025.

Education

Ph.D. Electrical Engineering and Applied Computing

2011-2016

Federal University of Pará with Split-time period - University of Kent

M.Sc in Electrical Engineering and Applied Computing

2010-2011

Federal University of Pará

Certifications

MBA Essentials

2022

London School of Economics and Political Science, LSE, United Kingdom

Advanced International Certificate Course on IP Asset Management for Business

2020

Korea Advanced Institute of Science and Technology, KAIST, South Korea

Selected Publications

Justino, A. F.; Jacob, A. F. L. Lobato, F.. Comparative Study of BERT Models for Semantic Retrieval of Brazilian Legal Precedents. *Symposium on Knowledge Discovery, Mining and Learning* (2025) 

Araujo, G. S.; Santana, E. E. C.; Jacob, A. F. L.; Lobato, F.. The Artificial Intelligence Integration in the Brazilian Legal Sector: A Systematic Review. *Simpósio Brasileiro de Sistemas de Informação* (2025) 

Golo, M. P. S.; Medeiros Jr., J. G. B.; Vilas Boas, G. S.; Lobato, F.; Silva, D. F.; Marcacini, R. M.. Large Language Models Ensemble for Biochemical Properties Detection in Scientific Articles *Second International Biochemical Knowledge Extraction Challenge (BiKE)* (2025) 

Leite, J. B. P.; Nhoatto, F. P.; Jacob Junior, A. F. L.; Hermida, R. S.; Lobato, F.. Computational Tools for Neuronal Morphometric Analysis: A Systematic Search and Review *Neuroinformatics* (2024) 

Jacob, A. F. L.; Do Carmo, F.; De Santana, A.; Santana, E.; Lobato, F.. Evolimp: Multiple Imputation of Multi-label Classification data with a genetic algorithm *PLoS One* (2024) 

do Carmo, F. A.; Silva Jr., J. L. F.; Rossi, R. G.; Lobato, Fábio. *Text representations for lyric-based identification of musical subgenres.* *IEEE Latin America Transactions* (2023) 

Akihito, Victor; Lobato, Fábio; Marcacini, Ricardo. *Weak Supervision for Question and Answering Sentiment Analysis* *International Conference on Machine Learning and Applications (ICMLA)*(2023) 

Goncalves, W. G. E.; Santos, M. H. P.; Brito, L. M.; Palheta, H. G. A.; Lobato, F.; Demachki, S.; Ribeiros Dos Santos, Â. K.; Araujo, G.. DeepHP: a new gastric mucosa histopathology dataset for Helicobacter pylori infection diagnosis. *International Journal of Molecular Sciences* (2022) 

Pedroso, P; Lobato, F.; SA, E; Jacob Jr. A. F. L.. Handling out of vocabulary words at the semantical level using recurrent neural networks. *21st IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* (2022) 

Relevant information

- Published more than 100 papers in high-level journals, conferences, and book chapters
- Tech Lead for an R&D&I project developing custom language models for the Justice Court, overseeing the creation and annotation of domain-specific datasets
- Experienced in funding applications and R&D,I project management, has a distinguished productivity scholarship from the Brazilian National Council for Scientific and Technological Development (CNPq)
- Demonstrated collaborative leadership across multicultural, cross-disciplinary teams, working closely with diverse subject matter experts—including Judges, Pathologists, Geneticists, and Engineers, to translate complex domain knowledge into robust, ethically-sound AI and NLP solutions