

Warning: [2026-06-04 18:19] this document is a print-out of the Ciência-iul web portal and was automatically generated at the labeled date. The document has a mere informational purpose and represents the information contained on Ciência_Iscte at that date.

Fernando Manuel Marques Batista

Professor Associado (com Agregação)

Department of Information Science and Technology (ISTA)



Contacts

| | |
|------------------|-------------------------------|
| E-mail | fernando.batista@iscte-iul.pt |
| Office | D6.21 |
| Telephone | 217650569 (Ext: 220700) |
| Post Box | 151 |

Curriculum

Fernando Batista received his PhD (2011) in Computer Science and Engineering from Instituto Superior Técnico (IST). He is currently Associate Professor with Aggregation at Iscte - University Institute of Lisbon, and integrated researcher at INESC-ID, Lisbon. He is the Executive Coordinator of the Human Language Technologies Scientific Area at INESC-ID (2021-), and member of the Supervisory Board of CVTT - Iscte-Conhecimento e Inovação (since Apr.2021). He was President of the Pedagogical Council of ISCTE-IUL (2017-2019), member of the Permanent Committee of the Pedagogical Council of ISCTE-IUL (2015-2017), and member of the Pedagogical Committee of the ISTA School (2015-2017). He was the scientific coordinator of AppRecommender (2019-2021), has coordinated the INESC-ID team in the SpeDial project (2014-2015), and participated in several other European and National projects. His current research focuses on spoken and written Natural Language Processing, Machine Learning, and Text Mining for social media. He has been member of the organisation team of the LxMLS - Lisbon Machine Learning Summer School (since 2016), and was also member of the LxMLS technical staff between (2011-2015). He has participated in the organisation of several international scientific events. He is Senior Member of the IEEE (since 2016), and member of the ISCA Speech (since 2008).

Research Interests

Natural Language Processing

| |
|-------------------|
| Text Mining |
| Machine Learning |
| Speech Processing |

Academic Qualifications

| University/Institution | Type | Degree | Period |
|---|-------------|---|--------|
| ISCTE-Instituto Universitario de Lisboa | Aggregation | Inteligência Artificial | 2025 |
| Instituto Superior Técnico (IST) | PhD | Engenharia Informática | 2011 |
| Instituto Superior Técnico (IST) | M.Sc. | Engenharia Electrotécnica e de Computadores | 2003 |
| Universidade da Beira Interior | Licenciate | Matemática/Informática ramo de Informática | 1997 |

Teaching Activities

| Teaching Year | Sem. | Course Name | Degree(s) | Coord |
|---------------|------|--|--|-------|
| 2026/2027 | 2º | Text Mining for Data Science | Master Degree in Data Science; | Yes |
| 2026/2027 | 2º | Applied Final Project in Data Science | Bachelor Degree in Data Science; | No |
| 2026/2027 | 1º | Data in Science, Bussiness and Society | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | Yes |
| 2026/2027 | 1º | Applied Project in Data Science | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | No |
| 2026/2027 | 1º | Advanced Machine Learning | Master Degree in Artificial Intelligence; | Yes |
| 2025/2026 | 2º | Text Mining for Data Science | Master Degree in Data Science; | Yes |
| 2025/2026 | 2º | Applied Final Project in Data Science | Bachelor Degree in Data Science; | No |
| 2025/2026 | 2º | NLP & Transformers | Other programme in Applied in AI for Tech Leaders; | Yes |
| 2025/2026 | 1º | Data in Science, Bussiness and Society | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | Yes |
| 2025/2026 | 1º | Applied Project in Data Science | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | No |

| | | | | |
|-----------|----|---|--|-----|
| 2025/2026 | 1° | Disruptive Technologies | Master Degree in Digital Technologies for Business; | No |
| 2025/2026 | 1° | Advanced Machine Learning | Master Degree in Artificial Intelligence; | Yes |
| 2024/2025 | 2° | Computational Language Processing | | No |
| 2024/2025 | 2° | Text Mining for Data Science | Master Degree in Data Science; | Yes |
| 2024/2025 | 2° | Applied Final Project in Data Science | Bachelor Degree in Data Science (PL); | No |
| 2024/2025 | 2° | Computing for Economics and Business | | Yes |
| 2024/2025 | 1° | Text Mining | | No |
| 2024/2025 | 1° | Data in Science, Bussiness and Society | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | Yes |
| 2024/2025 | 1° | Programming | | Yes |
| 2024/2025 | 1° | Applied Project in Data Science | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | No |
| 2024/2025 | 1° | Disruptive Technologies | Master Degree in Digital Technologies for Business; | No |
| 2024/2025 | 1° | Computer Vision and Introduction to Language Models | | No |
| 2023/2024 | 2° | Text Mining for Data Science | Master Degree in Data Science; | Yes |
| 2023/2024 | 2° | Computing for Economics and Business | | Yes |
| 2023/2024 | 1° | Computational Language Processing | Institutional Degree in Escola de Tecnologias e Arquitetura; | No |
| 2023/2024 | 1° | Text Mining | Master Degree in Integrated Business Intelligence Systems; | No |
| 2023/2024 | 1° | Data in Science, Bussiness and Society | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | Yes |
| 2023/2024 | 1° | Programming | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | Yes |
| 2023/2024 | 1° | Applied Project in Data Science | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | Yes |
| 2022/2023 | 2° | Text Mining for Data Science | Master Degree in Data Science; | Yes |
| 2022/2023 | 2° | Knowledge Discovery and Extraction From Data | Bachelor Degree in Computer Science and Business Management; | No |

| | | | | |
|-----------|----|--|--|-----|
| 2022/2023 | 2° | Computing for Economics and Business | | Yes |
| 2022/2023 | 1° | Computational Language Processing | | No |
| 2022/2023 | 1° | Text Mining | Master Degree in Integrated Business Intelligence Systems; | No |
| 2022/2023 | 1° | Programming | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | Yes |
| 2022/2023 | 1° | Applied Project in Data Science | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | Yes |
| 2021/2022 | 2° | Text Mining for Data Science | Master Degree in Data Science; | Yes |
| 2021/2022 | 2° | Knowledge Discovery and Extraction From Data | Bachelor Degree in Computer Science and Business Management; | No |
| 2021/2022 | 2° | Computing for Economics and Business | Institutional Degree in Escola de Tecnologias e Arquitetura; | Yes |
| 2021/2022 | 1° | Computational Language Processing | Institutional Degree in Escola de Tecnologias e Arquitetura; | No |
| 2021/2022 | 1° | Text Mining | Master Degree in Integrated Business Intelligence Systems; | No |
| 2021/2022 | 1° | Programming | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | Yes |
| 2021/2022 | 1° | Applied Project in Data Science | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | Yes |
| 2020/2021 | 2° | Text Mining for Data Science | Master Degree in Data Science; | Yes |
| 2020/2021 | 2° | Computing for Economics and Business | Institutional Degree in Escola de Tecnologias e Arquitetura; | Yes |
| 2020/2021 | 1° | Computational Language Processing | Institutional Degree in Escola de Tecnologias e Arquitetura; | Yes |
| 2020/2021 | 1° | Text Mining | Master Degree in Integrated Business Intelligence Systems; | Yes |
| 2020/2021 | 1° | Programming | Bachelor Degree in Data Science (PL); Bachelor Degree in Data Science; | Yes |
| 2018/2019 | 2° | Probabilistic Learning for Natural Language Processing | | Yes |
| 2018/2019 | 1° | Computational Language Processing | | Yes |
| 2018/2019 | 1° | Text Mining | Master Degree in Integrated Business Intelligence Systems; | No |

| | | | | |
|-----------|----|--|---|-----|
| 2018/2019 | 1° | Operating Systems | Bachelor Degree in Computer Engineering; Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | Yes |
| 2017/2018 | 2° | Probabilistic Learning for Natural Language Processing | | No |
| 2017/2018 | 2° | Text Mining | | Yes |
| 2017/2018 | 1° | Open Source Operating System | | Yes |
| 2017/2018 | 1° | Computational Language Processing | | No |
| 2017/2018 | 1° | Operating Systems | Bachelor Degree in Computer Engineering; Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | Yes |
| 2016/2017 | 2° | Probabilistic Learning for Natural Language Processing | | No |
| 2016/2017 | 2° | Text Mining | | No |
| 2016/2017 | 1° | Computational Language Processing | | Yes |
| 2016/2017 | 1° | Operating Systems | Bachelor Degree in Computer Engineering; Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | Yes |
| 2016/2017 | 1° | Operating Systems | Bachelor Degree in Computer Engineering; Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | Yes |
| 2015/2016 | 1° | Open Source Operating System | | No |
| 2015/2016 | 1° | Computational Language Processing | | No |
| 2015/2016 | 1° | Probabilistic Learning for Natural Language Processing | | Yes |
| 2015/2016 | 1° | Operating Systems | Bachelor Degree in Computer Science and Business Management (PL); Bachelor Degree in Computer Engineering (PL); Bachelor Degree in Computer Engineering; Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | Yes |

| | | | | |
|-----------|----|--|---|-----|
| 2015/2016 | 1° | Operating Systems | Bachelor Degree in Computer Science and Business Management (PL); Bachelor Degree in Computer Engineering (PL); Bachelor Degree in Computer Engineering; Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | Yes |
| 2014/2015 | 1° | Open Source Operating System | | Yes |
| 2014/2015 | 1° | Computational Language Processing | | Yes |
| 2014/2015 | 1° | Probabilistic Learning for Natural Language Processing | | No |
| 2014/2015 | 1° | Advanced Topics in Business Intelligence | | No |
| 2014/2015 | 1° | Operating Systems | Bachelor Degree in Computer Science and Business Management (PL); Bachelor Degree in Computer Engineering (PL); Bachelor Degree in Computer Engineering; Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | Yes |
| 2014/2015 | 1° | Operating Systems | Bachelor Degree in Computer Science and Business Management (PL); Bachelor Degree in Computer Engineering (PL); Bachelor Degree in Computer Engineering; Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | Yes |
| 2013/2014 | 1° | Open Source Operating System | | No |
| 2013/2014 | 1° | Computational Language Processing | | No |
| 2013/2014 | 1° | Operating Systems | Bachelor Degree in Computer Engineering; Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | Yes |
| 2013/2014 | 1° | Operating Systems | Bachelor Degree in Computer Engineering; Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | Yes |
| 2012/2013 | 1° | Open Source Operating System | | No |
| 2012/2013 | 1° | Operating Systems | Bachelor Degree in Computer Science and Business Management (PL); Bachelor Degree in Computer Engineering; Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | Yes |

| | | | | |
|-----------|----|------------------------------|--|----|
| 2011/2012 | 2º | Operating Systems | Bachelor Degree in Computer Science and Business Management (PL); Bachelor Degree in Computer Science and Business Management; Bachelor Degree in Telecommunications and Computer Engineering; | No |
| 2011/2012 | 1º | Open Source Operating System | | No |
| 2011/2012 | 1º | Computer Architecture | | No |
| 2010/2011 | 2º | Operating Systems | Bachelor Degree in Computer Science and Business Management (PL); Bachelor Degree in Computer Science and Business Management; | No |
| 2010/2011 | 1º | Computer Architecture | | No |
| 2009/2010 | 2º | Operating Systems | | No |
| 2007/2008 | 2º | Introduction to Programming | | No |
| 2007/2008 | 1º | Introduction to Programming | | No |

Supervisions

• Ph.D. Thesis

- Ongoing

| | Student Name | Title/Topic | Language | Status | Institution |
|---|----------------------------------|--|------------|------------|---|
| 1 | Eduardo Antonio Mello Freitas | "Language Complexity and Contextualized Translation". | English | Developing | Iscte |
| 2 | Maria Inês Bico | Aplicação de sistemas automáticos de processamento computacional da língua no Corpus de Textos Antigos | Portuguese | Developing | Faculdade de Letras da Universidade de Lisboa |
| 3 | Diogo Eduardo Ribeiro dos Santos | Project appraisal with pretrained transformer-based Large Language Models | English | Developing | Iscte |

- Concluded

| | Student Name | Title/Topic | Language | Institution | Concluding Year |
|---|---------------------------------------|--|----------|-------------|-----------------|
| 1 | Mariana Sofia Barreira Cavique Santos | Exploring the Perspectives of Hosts and Guests in Peer-to-peer accommodation using text mining | English | Iscte | 2024 |

• M.Sc. Dissertations

- Ongoing

| | Student Name | Title/Topic | Language | Status | Institution |
|---|-----------------------------------|--|----------|------------|-------------|
| 1 | Tomás Miguel Branco Vicente | Automatic Extraction, Classification, and Normalization of Entities from Clinical Notes | -- | Developing | Iscte |
| 2 | Ana Raquel de Matos Claudino | MatchMind - Optimizing Consultant-Project Allocation Through AI | -- | Developing | Iscte |
| 3 | Bruno Miguel Isidro Guarita | Automatic classification of drug-related health events from news feeds | -- | Developing | Iscte |
| 4 | Filipa Maria Barata Ramos | The Evolution of Housing Policy in Portuguese Electoral Programs: An Analysis Using NLP Techniques | -- | Developing | Iscte |
| 5 | Ricardo Araújo Galvão | Exploration of Artificial Intelligence methods based on graphs to represent and infer knowledge in the financial context | -- | Developing | Iscte |
| 6 | Renato Luís Ferreira Teixeira | Emotional Profile Identification and Toxicity Detection in Game Reviews | -- | Developing | Iscte |
| 7 | Diogo Alexandre Alonso de Freitas | Visual recognition of idiomatic expressions using multimodal models | -- | Developing | Iscte |
| 8 | Pedro António Ferra Moura | Ciberdúvidas: Conversational Agent for Answering Questions about the Portuguese Language | -- | Developing | Iscte |
| 9 | Filipa Franco | Recognizing Behavioral Elements in Team Interactions | -- | Developing | Iscte |

- Concluded

| | Student Name | Title/Topic | Language | Institution | Concluding Year |
|---|-----------------------------------|--|------------|-------------|-----------------|
| 1 | Humberto Carlos Pestana Fernandes | Automatic extraction and classification of textual content from periodicals in Portugal | Portuguese | Iscte | 2025 |
| 2 | Roberto Carlos Medina | Creating NLP resources to Cape Verdean Creole: Challenges and Solutions for Underrepresented Languages | English | Iscte | 2025 |
| 3 | Catarina Duarte Brito | Chatbot based on LLMs for Support and Development Teams | English | Iscte | 2025 |

| | | | | | |
|----|--|---|------------|-------|------|
| 4 | Guilherme Vaz Louro | Beyond Citation Counts: Evaluating Publication Impact Through Journal-Level Citation Metrics | English | Iscte | 2025 |
| 5 | Luís Miguel dos Santos Pereira | Unmasking Hidden Hate: Techniques for Detecting and Normalizing Obfuscated Offensive Language | English | Iscte | 2025 |
| 6 | Mauro Jorge Costa Cardoso | Unveiling the Far-Right's Political Ascent in Portugal: An NLP and AI Approach. | English | Iscte | 2025 |
| 7 | Jaciara Aly Ferreira | Unveiling Key Hotel Satisfaction Factors in Maputo Using Guest Reviews | English | Iscte | 2024 |
| 8 | Diana Carolina da Silva Monteiro Guimarães | The Circus Arts in the Era of Artificial Intelligence: A study of the current and future usage of A.I. in Circus Arts | English | Iscte | 2024 |
| 9 | Margarida Paias Moura | Predicting Success in Latin-Language Cinema: A Machine Learning Approach to Identify Key Factors in Film Performance | English | Iscte | 2024 |
| 10 | Gil Antunes Silva Nogueira Ramos | Automatic Hate Speech Detection in Portuguese Social Media | English | Iscte | 2024 |
| 11 | Marta Gameiro Francisco Moreno Pereira | From Regular Expressions to Language Models: A Text-Based Approach for Invoice Information Extraction | English | Iscte | 2024 |
| 12 | Inês Filipa Correia Morais | Exploring Feelings and Emotions in Dialogues: Analysis of Contact Center Interactions | Portuguese | Iscte | 2024 |
| 13 | Tiago Augusto Mendes Teixeira Barbosa | Melodies and Data: An Analytical Perspective on Predicting Human Behaviour in Music Learning. | Portuguese | Iscte | 2024 |
| 14 | Pedro Afonso Marques D'Oliveira | Navigating the Mobile App Galaxy: Harnessing Textual Metadata for App Categorization | English | Iscte | 2024 |
| 15 | Samuel Filipe Agostinho dos Santos | Optical Character Recognition, Text Analysis and Key Information Extraction | English | Iscte | 2024 |
| 16 | Mariana Patrícia Silva Ferreira | A Text Mining Approach to Portuguese Terroir: Analysing Online Wine Reviews | English | Iscte | 2024 |
| 17 | Filipa Camarate Gonçalves | Impact of Covid-19 Pandemic on the Dynamics of Accommodation Sharing: A Hosts Perspective | English | Iscte | 2024 |
| 18 | Matilde Lucena Leite da Silva | The Impacts of Cancel Culture on Brands and Celebrities | English | Iscte | 2024 |

| | | | | | |
|----|---|---|------------|--|------|
| 19 | Joana Maria Freire Fernandes | Optimization of Healthcare Practices: Integration of Big Data and Artificial Intelligence in Medical Diagnosis | Portuguese | Iscte | 2024 |
| 20 | Alexandre Costa Pereira | The use of Machine Learning and Text Mining in Stock Market Forecasting. | Portuguese | Iscte | 2023 |
| 21 | Miguel Alexandre da Silva Sarmiento Falco Pereira | Assessing NER Tools for Dialogue Data Anonymization | English | Iscte | 2023 |
| 22 | David Miguel Nunes Rodrigues | Automatic Detection of Missing Information in the Indexing of Scientific Publications | English | Iscte | 2023 |
| 23 | Alexandre Filipe Magalhães Cancelinha | Coding by voice in a projectional editor | English | Iscte | 2023 |
| 24 | Raquel Bento Santos | Semi-Automatic Selection and Annotation of Hate Speech from Social Media | English | Instituto Superior Técnico, Universidade de Lisboa | 2022 |
| 25 | Jana Luisa Teixeira Lage | Influencers, are they responsible for Bitcoin's volatility? Transfer Entropy and Granger causality in prol of an answer | English | Iscte | 2022 |
| 26 | Bruno Alexandre Mateus Francisco | Analysis of the tourists behavior in Lisbon using data from a Mobile Operator | English | Iscte | 2022 |
| 27 | Inês Mendes dos Santos Rodrigues de Carvalho | Identification of Risk Factors in the Severity of Road Accidents | Portuguese | Iscte | 2022 |
| 28 | Marcelo Custodio Mathias | Humor and Offense Speech Classification and scoring using Natural Language Processing | Portuguese | Iscte | 2022 |
| 29 | Érica Sofia Palmeirim Santos Rosa | What is People Talking About in your Region? Applying a Topic Modelling Approach to Portuguese Geolocated Tweets | English | Iscte | 2022 |
| 30 | Daniela Cristina da Silva Carlos | Automatic detection of depression signs in social networks in the context of covid-19 | Portuguese | Iscte | 2021 |
| 31 | Nuno Miguel Soares Fialho de Carvalho | Student Data Prediction | English | Iscte | 2021 |
| 32 | Cátia Daniela Lopes Tavares | Sentiment analysis to predict the Portuguese economic sentiment based on economic news | English | Iscte | 2021 |
| 33 | Appio Indiano do Brazil Americano Neto | Automatic fraud prediction in financial transactions | Portuguese | Iscte | 2021 |
| 34 | Tiago Filipe Pardal de Almeida | Towards Cyberbullying Detection on Social Media | English | Iscte | 2021 |
| 35 | Soraia Alexandra Cardoso Filipe | Classification of emotions in social networks | Portuguese | Iscte | 2021 |

| | | | | | |
|----|--------------------------------------|--|------------|--|------|
| 36 | André Filipe Lopes Maia | Assessing the Emotional Impact of Video using Machine Learning Techniques | English | Iscte | 2020 |
| 37 | Rosária Patrícia Firmino Bunga | Videogame-oriented Recommendation System | Portuguese | Iscte | 2020 |
| 38 | Afonso do Carmo Marques Mendes Pinto | Detection of emerging words in Portuguese tweets and analysis of their path in social networks | Portuguese | Iscte | 2020 |
| 39 | Nuno Alexandre Mestre Barradas | A book-oriented Chatbot | English | Iscte | 2020 |
| 40 | Miguel de Guerra Narciso | Cryptocurrency Analysis based on User-Generated Social Media Content | English | Iscte | 2020 |
| 41 | Patrícia de Sousa dos Santos | Detecting Psychological Sentiments in Users from Social Networks | English | Iscte | 2020 |
| 42 | João Filipe Carriço Rodrigues | Fake News Classification in European Portuguese Language | English | Iscte | 2020 |
| 43 | Marco Filipe Madeira Felgueiras | Multilabel classification of unstructured data using Crunchbase | English | Iscte | 2020 |
| 44 | Joana de Azinhaes Horta | eWOM for public institutions: application to the case of the Portuguese Army | Portuguese | Iscte | 2020 |
| 45 | Sónia Alexandra Miranda Morais | Modeling reviews on online platforms about local accommodation: The Airbnb case | Portuguese | Iscte | 2019 |
| 46 | João Nuno Martins de Almeida | 2Gather4Health: Web Crawling and Indexing System Implementation | English | Instituto Superior Técnico, Universidade de Lisboa | 2018 |
| 47 | Diana Von-Haff Lopes Teixeira | Spatio-Temporal Distribution Analysis of Brand Interest in Social Networks | English | Iscte | 2018 |
| 48 | Tiago João Aires Soares | Text Mining from CURRICULA | Portuguese | Iscte | 2018 |
| 49 | Ricardo Afonso Gomes António | Análise de Sentimento Baseada em Aspetos | Portuguese | Iscte | 2017 |
| 50 | Fernando Manuel Dias Rebelo | Análise de Sentimento em Microblogues com base em Cascatas de Classificação | Portuguese | Iscte | 2016 |
| 51 | Anzhela Zhusupova | Characterizing the Personality of Twitter Users Based on their Timeline Information | English | Iscte | 2016 |
| 52 | Luís Jorge Gregório Dias | Detecting Violent Excerpts in Movies using Audio and Video Features | English | Iscte | 2016 |
| 53 | Marco Paulo Fernandes Vicente | "Detecting portuguese and english twitter user's gender" | English | Iscte | 2015 |

| | | | | | |
|----|--|---|------------|--|------|
| 54 | Gaspar Manuel Rocha Brogueira | Sistema Inteligente de Recolha e Armazenamento de Informação Proveniente do Twitter | Portuguese | Iscte | 2015 |
| 55 | Hugo Hermógenes Lopes da Costa Rosa | Topic Detection within Public Social Networks | English | Instituto Superior Técnico, Universidade de Lisboa | 2014 |
| 56 | Henrique Rodrigues Barbosa de Medeiros | Automatic detection of disfluences in a corpus of univeristy lectures | English | Iscte | 2014 |

• M.Sc. Final Projects

- Concluded

| | Student Name | Title/Topic | Language | Institution | Concluding Year |
|---|------------------------------------|---|------------|-------------|-----------------|
| 1 | Nelson Carvalho de Jesus | From Legal Jargon to Generative Artificial Intelligence: The Legal System in the Age of the Transformer | Portuguese | Iscte | 2025 |
| 2 | Pedro Miguel Delgado do Nascimento | Automatic Transcription System of Parliamentary Debates in the Context of the Portuguese Parliament | English | Iscte | 2024 |

Total Citations

| | |
|-----------------|-----|
| Web of Science® | 557 |
| Scopus | 837 |

Publications

• Scientific Journals

- Scientific journal paper

| | |
|---|--|
| 1 | Rodrigues, D., Lopes, A. & Batista, F. (N/A). Detecting incoherent citation data among three bibliometric platforms: OpenAlex, Scopus and Web of Science. Journal of Information Science. N/A - Times Cited Web of Science®: 5 - Times Cited Scopus: 4 - Times Cited Google Scholar: 6 |
| 2 | Taborda, B., Almeida, A. M. de., Dias, J. C., Batista, F. & Ribeiro, R. (2025). SA-MAIS: Hybrid automatic sentiment analyser for stock market. Journal of Information Science. 51 (6), 1443-1456 - Times Cited Web of Science®: 1 - Times Cited Scopus: 1 - Times Cited Google Scholar: 2 |

| | |
|----|---|
| 3 | <p>Guerra, R., Carvalho, P., Marques, C., Carmona, M., Sarroeira, R., Batista, F....Silva, C. (2025). Unpacking online hate speech in Portuguese social media: A social-psychological and linguistic-discursive approach. <i>Humanities and Social Sciences Communications</i>. 12</p> <p>- Times Cited Web of Science®: 1</p> <p>- Times Cited Google Scholar: 1</p> |
| 4 | <p>Cavique, M., Ribeiro, R., Batista, F. & Correia, A. (2025). From host's descriptions to guests' reviews: Semantic similarities. <i>Annals of Tourism Research Empirical Insights</i>. 6 (2)</p> |
| 5 | <p>Bico, M. I., Carneira, E., Baptista, J. & Batista, F. (2025). Evidências do Português Médio no Corpus de Textos Antigos. <i>Estudos de Lingüística Galega</i>. 17</p> |
| 6 | <p>Sanchez-Gomez, J. M., Batista, F., Vega-Rodríguez, M. A. & Pérez, C. J. (2025). A transformer-based deep learning approach for detecting online hate speech in Spanish. <i>Applied Soft Computing</i>. 187</p> |
| 7 | <p>Bico, M. I., Baptista, J., Batista, F. & Carneira, E. (2024). Enriching Portuguese medieval texts with named entity recognition. <i>International Journal of Humanities and Arts Computing</i>. 18 (1), 109-124</p> |
| 8 | <p>Carvalho, P., Caled, D., Silva, C., Batista, F. & Ribeiro, R. (2024). The expression of hate speech against Afro-descendant, Roma, and LGBTQ+ communities in YouTube comments. <i>Journal of Language Aggression and Conflict</i> . 12 (2), 171-206</p> <p>- Times Cited Web of Science®: 21</p> <p>- Times Cited Scopus: 18</p> <p>- Times Cited Google Scholar: 38</p> |
| 9 | <p>Nascimento, P., Ferreira, J. C. & Batista, F. (2024). Automatic transcription system for parliamentary debates in the context of assembly of the republic of Portugal. <i>International Journal of Speech Technology</i>. 27 (3), 613-635</p> <p>- Times Cited Web of Science®: 1</p> <p>- Times Cited Scopus: 2</p> <p>- Times Cited Google Scholar: 4</p> |
| 10 | <p>Ramos, G., Batista, F., Ribeiro, R., Fialho, P., Moro, S., Fonseca, A....Silva, C. (2024). Leveraging transfer learning for hate speech detection in Portuguese social media posts. <i>IEEE Access</i>. 12, 101374-101389</p> <p>- Times Cited Web of Science®: 9</p> <p>- Times Cited Scopus: 12</p> <p>- Times Cited Google Scholar: 20</p> |
| 11 | <p>Fonseca, A., Pontes, C., Moro, S., Batista, F., Ribeiro, R., Guerra, R....Silva, C. (2024). Analyzing hate speech dynamics on Twitter/X: Insights from conversational data and the impact of user interaction patterns. <i>Heliyon</i>. 10 (11)</p> <p>- Times Cited Web of Science®: 6</p> <p>- Times Cited Scopus: 6</p> <p>- Times Cited Google Scholar: 19</p> |
| 12 | <p>Peixoto, A. R., Almeida, A. de., Antonio, N., Batista, F., Ribeiro, R. & Cardoso, E. (2023). Unlocking the power of Twitter communities for startups. <i>Applied Network Science</i>. 8</p> <p>- Times Cited Web of Science®: 2</p> <p>- Times Cited Google Scholar: 7</p> |
| 13 | <p>Peixoto, A., de Almeida, A., Antonio, N., Batista, F. & Ribeiro, R. (2023). Diachronic profile of startup companies through social media. <i>Social Network Analysis and Mining</i>. 13</p> <p>- Times Cited Web of Science®: 6</p> <p>- Times Cited Scopus: 6</p> <p>- Times Cited Google Scholar: 12</p> |

| | |
|----|--|
| 14 | <p>Coelho, J., Mano, D., Paula, B., Coutinho, C., Oliveira, J., Ribeiro, R....Batista, F. (2023). Semantic similarity for mobile application recommendation under scarce user data. <i>Engineering Applications of Artificial Intelligence</i>. 121</p> <p>- Times Cited Web of Science®: 7 - Times Cited Scopus: 6 - Times Cited Google Scholar: 9</p> |
| 15 | <p>Santos, M., Correia, A., Ribeiro, R. & Batista, F. (2022). What are Airbnb hosts advertising? A longitudinal essay in Lisbon. <i>Consumer Behavior in Tourism and Hospitality</i>. 17 (3), 312-325</p> <p>- Times Cited Web of Science®: 5 - Times Cited Scopus: 6 - Times Cited Google Scholar: 8</p> |
| 16 | <p>Santos, M., Ribeiro, R., Batista, F. & Correia, A. (2022). Examining Airbnb guest satisfaction tendencies: A text mining approach. <i>Current Issues in Tourism</i>. 25 (22), 3607-3622</p> <p>- Times Cited Web of Science®: 20 - Times Cited Scopus: 17 - Times Cited Google Scholar: 25</p> |
| 17 | <p>Guerreiro, N. M., Rei, R. & Batista, F. (2021). Towards better subtitles: A multilingual approach for punctuation restoration of speech transcripts. <i>Expert Systems with Applications</i>. 186</p> <p>- Times Cited Web of Science®: 7 - Times Cited Scopus: 12 - Times Cited Google Scholar: 15</p> |
| 18 | <p>Azinhaes, J., Batista, F. & Ferreira, J. (2021). eWOM for public institutions: application to the case of the Portuguese Army. <i>Social Network Analysis and Mining</i>. 11 (1)</p> <p>- Times Cited Web of Science®: 5 - Times Cited Scopus: 7 - Times Cited Google Scholar: 8</p> |
| 19 | <p>Fernandes, E., Moro, S., Cortez, P., Batista, F. & Ribeiro, R. (2021). A data-driven approach to measure restaurant performance by combining online reviews with historical sales data. <i>International Journal of Hospitality Management</i>. 94</p> <p>- Times Cited Web of Science®: 39 - Times Cited Scopus: 39 - Times Cited Google Scholar: 76</p> |
| 20 | <p>Cabarrão, V., Batista, F., Moniz, H., Trancoso, I. & Mata, A. I. (2019). Adaptação acústico-prosódica local em Português Europeu. <i>Revista da Associação Portuguesa de Linguística</i>. 5, 79-94</p> |
| 21 | <p>Solera-Ureña, R., Moniz, H., Batista, F., Cabarrão, V., Pompili, A., Astudillo, R....Trancoso, I. (2019). Uma abordagem de aprendizagem semissupervisionada para a classificação automática de personalidade baseada em pistas acústico-prosódicas. <i>Revista da Associação Portuguesa de Linguística</i>. 5, 348-364</p> |
| 22 | <p>Moro, S., Batista, F., Rita, P., Oliveira, C. & Ribeiro, R. (2019). Are the States United? An analysis of US hotels' offers through TripAdvisor's eyes. <i>Journal of Hospitality and Tourism Research</i>. 43 (7), 1112-1129</p> <p>- Times Cited Web of Science®: 15 - Times Cited Scopus: 13 - Times Cited Google Scholar: 25</p> |
| 23 | <p>Cabarrão, V., Moniz, H., Batista, F., Ferreira, J., Trancoso, I. & Mata, A. I. (2018). Cross-domain analysis of discourse markers in European Portuguese. <i>Dialogue and Discourse</i>. 9 (1), 79-106</p> <p>- Times Cited Web of Science®: 1 - Times Cited Scopus: 4 - Times Cited Google Scholar: 15</p> |

| | |
|----|--|
| 24 | Moro, S., Rita, P., Oliveira, C., Batista, F. & Ribeiro, R. (2018). Leveraging national tourist offices through data analytics. <i>International Journal of Culture, Tourism, and Hospitality Research</i> . 12 (4), 420-426 - Times Cited Scopus: 7 - Times Cited Google Scholar: 16 |
| 25 | Cabarrão, V., Moniz, H., Batista, F., Trancoso, I. & Mata, A. I. (2018). Adaptação acústico-prosódica entre falantes. <i>Revista da Associação Portuguesa de Linguística</i> . 4, 18-33 - Times Cited Google Scholar: 1 |
| 26 | Antonio, N., de Almeida, A. M., Nunes, L., Batista, F. & Ribeiro, R. (2018). Hotel online reviews: creating a multi-source aggregated index. <i>International Journal of Contemporary Hospitality Management</i> . 30 (12), 3574-3591 - Times Cited Web of Science®: 33 - Times Cited Scopus: 37 - Times Cited Google Scholar: 69 |
| 27 | Antonio, N., de Almeida, A., Nunes, L., Batista, F. & Ribeiro, R. (2018). Hotel online reviews: different languages, different opinions. <i>Information Technology and Tourism</i> . 18 (1-4), 157-185 - Times Cited Web of Science®: 34 - Times Cited Scopus: 39 - Times Cited Google Scholar: 75 |
| 28 | Carvalho, J. P., Rosa, H., Brogueira, G. & Batista, F. (2017). MISNIS: an intelligent platform for Twitter topic mining. <i>Expert Systems with Applications</i> . 89, 374 - 388 - Times Cited Web of Science®: 36 - Times Cited Scopus: 41 - Times Cited Google Scholar: 64 |
| 29 | Rebello, C., Pereira, I., Rosa, H., Batista, F. & Carvalho, J. P. (2017). Twitter: de espaço de mobilização a espaço de comentário. O caso do meet do Vasco da Gama. <i>Observatorio (OBS*)</i> . 11 (4), 19-41 - Times Cited Google Scholar: 5 |
| 30 | Brogueira, G., Batista, F. & Carvalho, J. P. (2017). A smart system for twitter corpus collection, management and visualization. <i>International Journal of Technology and Human Interaction</i> . 13 (3), 13-32 - Times Cited Web of Science®: 3 - Times Cited Scopus: 4 - Times Cited Google Scholar: 7 |
| 31 | Cabarrão, V., Moniz, H., Ferreira, J., Batista, F., Trancoso, I., Mata, Ana I....Curto, S. (2016). Classificação prosódica de marcadores discursivos. <i>Revista da Associação Portuguesa de Linguística</i> . 2, 69-75 - Times Cited Web of Science®: 1 - Times Cited Google Scholar: 1 |
| 32 | Brogueira, G., Batista, F. & Carvalho, J. P. (2016). Using geolocated tweets for characterization of Twitter in Portugal and the Portuguese administrative regions. <i>Social Network Analysis and Mining</i> . 6 (1) - Times Cited Web of Science®: 6 - Times Cited Scopus: 6 - Times Cited Google Scholar: 15 |
| 33 | Moniz, H., Batista, F., Mata, A. I. & Trancoso, I. (2014). Speaking style effects in the production of disfluencies. <i>Speech Communication</i> . 65, 20-35 - Times Cited Web of Science®: 25 - Times Cited Scopus: 20 - Times Cited Google Scholar: 59 |

| | |
|----|--|
| 34 | Barreiro, A., Ling, W, Coheur, L., Batista, F. & Trancoso, I. (2014). Projetos sobre tradução automática do português no laboratório de sistemas de língua falada do INESC-ID. <i>Linguamática</i> . 6 (2), 75-85 - Times Cited Google Scholar: 6 |
| 35 | Mata, Ana I., Moniz, H. & Batista, F. (2014). Comparing phrase-final patterns across speech styles and groups in European Portuguese. <i>Nouveaux Cahiers de Linguistique française</i> . 31 , 171-176 - Times Cited Google Scholar: 2 |
| 36 | Batista, F. & Ribeiro, R. (2013). Sentiment analysis and topic classification based on binary maximum entropy classifiers. <i>Procesamiento del Lenguaje Natural</i> . 50, 77-84 - Times Cited Scopus: 25 - Times Cited Google Scholar: 43 |
| 37 | Batista, F., Moniz, H., Trancoso, I. & Mamede, N. (2012). Bilingual Experiments on Automatic Recovery of Capitalization and Punctuation of Automatic Speech Transcripts. <i>IEEE Transactions on Audio, Speech, and Language Processing</i> . 20 (2), 474-485 - Times Cited Web of Science®: 42 - Times Cited Scopus: 51 - Times Cited Google Scholar: 86 |
| 38 | Batista, F., Moniz, H., Trancoso, I., Mamede, N. & Mata, A. I. (2012). Extending automatic transcripts in a unified data representation towards a prosodic-based metadata annotation and evaluation. <i>Journal of Speech Sciences</i> . 2 (2), 115-138 - Times Cited Google Scholar: 18 |
| 39 | Batista, F., Caseiro, D., Mamede, N. & Trancoso, I. (2008). Recovering capitalization and punctuation marks for automatic speech recognition: case study for Portuguese broadcast news. <i>Speech Communication</i> . 50 (10), 847-862 - Times Cited Web of Science®: 31 - Times Cited Scopus: 41 - Times Cited Google Scholar: 57 |

- Review article

| | |
|---|---|
| 1 | Ramos, G., Batista, F., Ribeiro, R., Fialho, P., Moro, S., Fonseca, A....Silva, C. (2024). A comprehensive review on automatic hate speech detection in the age of the transformer. <i>Social Network Analysis and Mining</i> . 14 (1) - Times Cited Web of Science®: 10 - Times Cited Scopus: 22 - Times Cited Google Scholar: 41 |
|---|---|

• Books and Book Chapters

- Book chapter

| | |
|---|---|
| 1 | Medina, R., Batista, F. & Lima, Estanislau (2025). CrioleSet: A New Corpus for Cape Verdean Creole, Towards Robust Machine Translation. In <i>Progress in Artificial Intelligence</i> . (pp. 326-338). - Times Cited Google Scholar: 1 |
| 2 | Bico, M. I., Baptista, J., Batista, F. & Cardeira, E. (2024). Recognizing and linking named entities in Portuguese medieval texts. In Sílvia Araújo; Micaela Aguiar; Liana Ermakova (Ed.), <i>Digital humanities looking at the world: Exploring innovative approaches and contributions to society</i> . (pp. 329-337). Cham: Springer Nature. |

| | |
|----|--|
| 3 | <p>Rebello-Arnold, I., Barreiro, A., Kuhn, T., Garcez, I. & Batista, F. (2021). Análise comparativa das edições portuguesa e brasileira da obra os livros que devoraram o meu pai, Afonso Cruz. In Rodrigo Santos de Oliveira, Michelle Vasconcelos Oliveira do Nascimento (Ed.), História, política e cultura no mundo lusófono. (pp. 119-131).: Editora LiberArs.</p> |
| 4 | <p>Vicente, M., Batista, F. & Carvalho, J. P. (2019). Gender detection of Twitter users based on multiple information sources. In Kóczy, László T. and Medina-Moreno, Jesús; Ramírez-Poussa, Eloísa (Ed.), Interactions Between Computational Intelligence and Mathematics Part 2. Studies in Computational Intelligence. (pp. 39-54). Cham: Springer International Publishing .</p> <p>- Times Cited Scopus: 17</p> <p>- Times Cited Google Scholar: 51</p> |
| 5 | <p>Rosa, H., Carvalho, J. P., Astudillo, R. & Batista, F. (2018). Page rank versus katz: is the centrality algorithm choice relevant to measure user influence in Twitter?. In Kóczy, László T.; Medina, Jesús (Ed.), Interactions Between Computational Intelligence and Mathematics. Studies in Computational Intelligence. Cham: Springer.</p> <p>- Times Cited Scopus: 8</p> <p>- Times Cited Google Scholar: 9</p> |
| 6 | <p>Moniz, H., Batista, F., Mata, Ana I. & Trancoso, I. (2016). Towards automatic language processing and intonational labeling in European Portuguese. In Meghan E. Armstrong and Nicholas C Henriksen and Maria del Mar Vanrell (Ed.), Intonational Grammar in Ibero-Romance: Approaches across linguistic subfields. (pp. 227-248). Amsterdam: John Benjamins Publishing Company.</p> <p>- Times Cited Google Scholar: 4</p> |
| 7 | <p>Mata, A. I., Moniz, H. & Batista, F. (2016). Stylistic variation in the intonation of European Portuguese teenagers and adults. In Meghan E. Armstrong and Nicholas C Henriksen and Maria del Mar Vanrell (Ed.), Intonational Grammar in Ibero-Romance: Approaches across linguistic subfields. (pp. 45-68). Amsterdam: John Benjamins Publishing Company.</p> <p>- Times Cited Google Scholar: 3</p> |
| 8 | <p>Moura, J., Batista, F., Cardoso, E. & Nunes, L. (2015). Intelligent management and efficient operation of big data. In Noor Zaman, Mohamed Elhassan Seliaman, Mohd Fadzil Hassan, Fausto Pedro Garcia Marquez (Ed.), Handbook of research on trends and future directions in big data and web intelligence. (pp. 105-129).: IGI Global.</p> <p>- Times Cited Google Scholar: 5</p> |
| 9 | <p>M. Julião, J. Silva, A. Aguiar, Moniz, H. & Batista, F. (2015). Speech Features for Discriminating Stress Using Branch and Bound Wrapper Search. In Sierra-Rodríguez, José-Luis and Leal, José-Paulo and Simões, Alberto (Ed.), Languages, Applications and Technologies. (pp. 3-14).: Springer International Publishing.</p> <p>- Times Cited Web of Science®: 3</p> <p>- Times Cited Scopus: 3</p> <p>- Times Cited Google Scholar: 12</p> |
| 10 | <p>Moniz, H., Batista, F., Trancoso, I. & Mata, Ana I. (2011). Analysis of interrogatives in different domains. In Third COST 2102 International Training School (Ed.), Proceedings of the Third COST 2102 international training school conference on Toward autonomous, adaptive, and context-aware multimodal interfaces: theoretical and practical issues. (pp. 136-148). Caserta, Italy: Springer Berlin / Heidelberg.</p> <p>- Times Cited Web of Science®: 5</p> <p>- Times Cited Scopus: 7</p> <p>- Times Cited Google Scholar: 13</p> |
| 11 | <p>Batista, F. & Mamede, N. (2004). 7 - SuSAna - Shallow Parsing With Preferences. In António Branco and Amália Mendes and Ricardo Ribeiro (Ed.), Language Technology for Portuguese: shallow processing tools and resources. Lisbon: Edições Colibri.</p> |

• **Conferences/Workshops and Talks**

- **Publication in conference proceedings**

| | |
|---|---|
| 1 | Medina, Roberto Carlos, Batista, F. & Lima, Estanislau (2025). CrioleSet: A New Corpus for Cape Verdean Creole, Towards Robust Machine Translation. In Valente de Oliveira, José and Leite, João and Rodrigues, João and Dias, João and Cardoso, Pedro (Ed.), Progress in Artificial Intelligence. (pp. 326-338). Cham: Springer Nature Switzerland. |
| 2 | Ramos, G., Batista, F., Ribeiro, R., Fialho, P., Moro, S., Fonseca, A....Silva, C. (2025). Bypassing the Nuances of Portuguese Covert Hate Speech through Contextual Analysis. In Progress in Artificial Intelligence. EPIA 2024. Lecture Notes in Computer Science, vol. 14969. (pp. 241-253). Viana do Castelo: Springer. - Times Cited Google Scholar: 1 |
| 3 | Cardoso, M., Ribeiro, E. & Batista, F. (2025). Portuguese Far-Right Discourse on Social Media: Insights from Topic Modeling. In 14th Symposium on Languages, Applications and Technologies (SLATE 2025). Faro |
| 4 | Pontes, C., Fonseca, A., Moro, S., Batista, F., Ribeiro, R., Marques, C....Guerra, R. (2025). Unveiling Patterns of Hate Speech in the Portuguese Sphere: A Social Network Analysis Approach. In Information Processing and Management of Uncertainty in Knowledge-Based Systems. (pp. 70-81). - Times Cited Google Scholar: 2 |
| 5 | Fialho, P., Ribeiro, R., Batista, F., Ramos, G., Fonseca, A., Moro, S....Silva, C. (2025). Counter Hate Speech Detection in Youtube Conversations. In Information Processing and Management of Uncertainty in Knowledge-Based Systems. (pp. 94-105). - Times Cited Google Scholar: 1 |
| 6 | Santos, A. L., Cancelinha, A. & Batista, F. (2024). Jasay: Towards voice commands in projectional editors. In Proceedings - 2024 1st IDE Workshop, IDE 2024. (pp. 30-34). Lisbon, Portugal: Association for Computing Machinery. - Times Cited Web of Science®: 1 - Times Cited Scopus: 1 - Times Cited Google Scholar: 1 |
| 7 | Bico, M. I., Carneira, E., Baptista, J. & Batista, F. (2024). Support verb constructions in medieval Portuguese: Evidence from the CTA corpus. In Pablo Gamallo, Daniela Claro, António Teixeira, Livy Real, Marcos Garcia, Hugo Gonçalo Oliveira, Raquel Amaro (Ed.), Proceedings of the 16th International Conference on Computational Processing of Portuguese: Vol.2. (pp. 118-129).: Association for Computational Linguistics. - Times Cited Google Scholar: 1 |
| 8 | Rodrigues, D., Lopes, A. L. & Batista, F. (2023). Web of Science citation gaps: An automatic approach to detect indexed but missing citations. In Simões, A., Berón, M. M., and Portela, F. (Ed.), 12th Symposium on Languages, Applications and Technologies (SLATE 2023). Porto: Schloss Dagstuhl -- Leibniz-Zentrum für Informatik. - Times Cited Web of Science®: 2 - Times Cited Scopus: 2 - Times Cited Google Scholar: 5 |
| 9 | Francisco, B., Ribeiro, R., Batista, F. & Ferreira, J. (2023). Analysis of the tourist's behavior in Lisbon using data from a mobile operator. In Martins, A. L., Ferreira, J. C., Kocian, A., and Tokkozhina, U. (Ed.), Intelligent Transport Systems. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. (pp. 3-22). Lisboa: Spinger. - Times Cited Google Scholar: 4 |

| | |
|----|--|
| 10 | <p>Vicente, M., Guarda, J. & Batista, F. (2022). Gutenbrain: An architecture for equipment technical attributes extraction from Piping & Instrumentation Diagrams. In Coenen, F., Fred, A., and Filipe, J. (Ed.), Proceedings of the 14th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management - KDIR. (pp. 204-211). Valletta, Malta: SCITEPRESS – Science and Technology Publications, Lda.</p> <p>- Times Cited Google Scholar: 2</p> |
| 11 | <p>Bico, M. I., Baptista, J., Batista, F. & Cardeira, E. (2022). Early experiments on automatic annotation of Portuguese medieval texts. In Silvello, G., Corcho, O., Manghi, P., Di Nunzio, G. M., Golub, K., Ferro, N., and Poggi, A. (Ed.), Linking theory and practice of digital libraries. Lecture Notes in Computer Science. (pp. 442-449). Padua: Springer International Publishing.</p> <p>- Times Cited Web of Science®: 2 - Times Cited Scopus: 3 - Times Cited Google Scholar: 5</p> |
| 12 | <p>Carvalho, P., Matos, B., Santos, R., Batista, F. & Ribeiro, R. (2022). Hate speech dynamics against African descent, Roma and LGBTQ+ communities in Portugal. In Calzolari, N., Béchet, F., Blache, P., Choukri, K., Cieri, C., Declerck, T., Goggi, S., Isahara, H., Maegaard, B., Mariani, J., Mazo, H., Odijk, J., and Piperidis, S. (Ed.), Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022). (pp. 2362-2370). Marseille: European Language Resources Association (ELRA).</p> <p>- Times Cited Scopus: 15 - Times Cited Google Scholar: 23</p> |
| 13 | <p>Matos, B. C., Santos, R. B., Carvalho, P., Ribeiro, R. & Batista, F. (2022). Comparing different approaches for detecting hate speech in online Portuguese comments. In Cordeiro, J., Pereira, M. J., Rodrigues, N. F., and Pais, S. (Ed.), OpenAccess Series in Informatics. Covilhã: Schloss Dagstuhl- Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing.</p> <p>- Times Cited Web of Science®: 2 - Times Cited Scopus: 3 - Times Cited Google Scholar: 10</p> |
| 14 | <p>Santos, R. B., Matos, B. C., Carvalho, P., Batista, F. & Ribeiro, R. (2022). Semi-supervised annotation of Portuguese hate speech across social media domains. In Cordeiro, J., Pereira, M. J., Rodrigues, N. F., and Pais, S. (Ed.), OpenAccess Series in Informatics. Covilhã: Schloss Dagstuhl- Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing.</p> <p>- Times Cited Web of Science®: 4 - Times Cited Scopus: 7 - Times Cited Google Scholar: 17</p> |
| 15 | <p>Paula, B., Coelho, J., Mano, D., Coutinho, C., Oliveira, J., Ribeiro, R....Batista, F. (2022). Collaborative filtering for mobile application recommendation with implicit feedback. In Morel, L., Dupont, L., and Camargo, M. (Ed.), 2022 IEEE 28th International Conference on Engineering, Technology and Innovation (ICE/ITMC) and 31st International Association For Management of Technology (IAMOT) Joint Conference. (pp. 1065 - 1073). Nancy, France: IEEE.</p> <p>- Times Cited Web of Science®: 1 - Times Cited Scopus: 3 - Times Cited Google Scholar: 5</p> |
| 16 | <p>Coelho, J., Neto, A., Tavares, M., Coutinho, C., Oliveira, J., Ribeiro, R....Batista, F. (2021). Transformer-based language models for semantic search and mobile applications retrieval. In Cucchiara, R., Fred, A., & Filipe, J. (Ed.), Proceedings of the 13th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. (pp. 225-232): SCITEPRESS – Science and Technology Publications, Lda.</p> <p>- Times Cited Web of Science®: 3 - Times Cited Google Scholar: 10</p> |

| | |
|----|--|
| 17 | <p>Bunga, R., Batista, F. & Ribeiro, R. (2021). From implicit preferences to ratings: Video games recommendation based on collaborative filtering. In Cucchiara, R., Fred, A., & Filipe, J. (Ed.), Proceedings of the 13th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. (pp. 209-216).: SCITEPRESS – Science and Technology Publications, Lda.</p> <p>- Times Cited Web of Science®: 3 - Times Cited Google Scholar: 12</p> |
| 18 | <p>Tavares, C., Ribeiro, R. & Batista, F. (2021). Sentiment analysis of Portuguese economic news. In Queirós, R., Pinto, M., Simões, A., Portela, F., & Pereira, M. J. (Ed.), 10th Symposium on Languages, Applications and Technologies (SLATE 2021). Vila do Conde/Póvoa de Varzim: Schloss Dagstuhl- Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing.</p> <p>- Times Cited Web of Science®: 2 - Times Cited Scopus: 2 - Times Cited Google Scholar: 10</p> |
| 19 | <p>Coelho, J., Neto, A., Tavares, M., Coutinho, C., Ribeiro, R. & Batista, F. (2021). Semantic search of mobile applications using word embeddings. In Queirós, R., Pinto, M., Simões, A., Portela, F., & Pereira, M. J. (Ed.), 10th Symposium on Languages, Applications and Technologies (SLATE 2021). Vila do Conde/Póvoa de Varzim: Schloss Dagstuhl- Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing.</p> <p>- Times Cited Web of Science®: 1 - Times Cited Scopus: 2 - Times Cited Google Scholar: 8</p> |
| 20 | <p>Rei, R., Batista, F., Guerreiro, N. M. & Coheur, L. (2021). Multilingual simultaneous sentence end and punctuation prediction. In Benites, F., Tuggener, D., Hurlimann, M., Cieliebak, M., & Vogel, M. (Ed.), Proceedings of the Swiss Text Analytics Conference 2021. Winterthur: CEUR-WS.</p> <p>- Times Cited Google Scholar: 2</p> |
| 21 | <p>Ribeiro, E., Ribeiro, R., Batista, F. & Oliveira, J. (2020). Using topic information to improve non-exact keyword-based search for mobile applications. In Lesot, Marie-Jeanne and Vieira, Susana and Reformat, Marek Z. and Carvalho, João Paulo and Wilbik, Anna and Bouchon-Meunier, Bernadette and Yager, Ronald R. (Ed.), Information processing and management of uncertainty in knowledge-based systems. (pp. 373-386).: Springer.</p> <p>- Times Cited Web of Science®: 2 - Times Cited Scopus: 4 - Times Cited Google Scholar: 6</p> |
| 22 | <p>Pinto, A., Moniz, H. & Batista, F. (2020). Detection of emerging words in Portuguese tweets. In Alberto Simões, Pedro Rangel Henriques, Ricardo Queirós (Ed.), 9th Symposium on Languages, Applications and Technologies (SLATE 2020). (pp. 3:1-3:10).: Schloss Dagstuhl--Leibniz-Zentrum für Informatik.</p> |
| 23 | <p>Filipe, S., Batista, F. & Ribeiro, R. (2020). Different lexicon-based approaches to emotion identification in Portuguese tweets. In Alberto Simões, Pedro Rangel Henriques, Ricardo Queirós (Ed.), 9th Symposium on Languages, Applications and Technologies (SLATE 2020). (pp. 12:1-12:8).: Schloss Dagstuhl--Leibniz-Zentrum für Informatik.</p> <p>- Times Cited Web of Science®: 1 - Times Cited Scopus: 1 - Times Cited Google Scholar: 3</p> |
| 24 | <p>Rodrigues, J., Ribeiro, R. & Batista, F. (2020). Towards the identification of fake news in Portuguese. In Alberto Simões, Pedro Rangel Henriques, Ricardo Queirós (Ed.), 9th Symposium on Languages, Applications and Technologies (SLATE 2020). (pp. 7:1-7:14).: Schloss Dagstuhl--Leibniz-Zentrum für Informatik.</p> |

| | |
|----|---|
| 25 | <p>Barreiro, A., Rebelo-Arnold, I., Batista, F., Garcez, I. & Kuhn, T. (2020). One book, two language varieties. In Quaresma, Paulo; Vieira, Renata; Alúcio, Sandra; Moniz, Helena; Batista, Fernando; Gonçalves, Teresa (Ed.), Computational Processing of the Portuguese Language .PROPOR 2020. Lecture notes in Computer Science. (pp. 379-389). Évora: Springer International Publishing.</p> <ul style="list-style-type: none"> - Times Cited Web of Science®: 2 - Times Cited Scopus: 3 - Times Cited Google Scholar: 5 |
| 26 | <p>Marco Felgueiras, Batista, F. & João P. Carvalho (2020). Creating classification models from textual descriptions of companies using crunchbase. In Lesot, Marie-Jeanne and Vieira, Susana and Reformat, Marek Z. and Carvalho, João Paulo and Wilbik, Anna and Bouchon-Meunier, Bernadette and Yager, Ronald R. (Ed.), Information Processing and Management of Uncertainty in Knowledge-Based Systems. (pp. 695-707). Lisboa: Springer International Publishing.</p> <ul style="list-style-type: none"> - Times Cited Web of Science®: 6 - Times Cited Scopus: 7 - Times Cited Google Scholar: 16 |
| 27 | <p>Ricardo Rei, Nuno Miguel Guerreiro & Batista, F. (2020). Automatic truecasing of video subtitles using BERT: a multilingual adaptable approach. In Lesot, Marie-Jeanne and Vieira, Susana and Reformat, Marek Z. and Carvalho, João Paulo and Wilbik, Anna and Bouchon-Meunier, Bernadette and Yager, Ronald R. (Ed.), Information Processing and Management of Uncertainty in Knowledge-Based Systems. (pp. 708-721): Springer International Publishing.</p> <ul style="list-style-type: none"> - Times Cited Web of Science®: 2 - Times Cited Scopus: 6 - Times Cited Google Scholar: 11 |
| 28 | <p>Cabarrão, V., Julião, M., Solera-Ureña, R., Moniz, H., Batista, F., Trancoso, I....Mata, Ana I. (2019). Affective analysis of customer service calls. In Botinis, A. (Ed.), Proceedings ExLing 2019: 10th International Conference of Experimental Linguistics. (pp. 37-40). Lisbon, Portugal: ExLing Society.</p> <ul style="list-style-type: none"> - Times Cited Google Scholar: 3 |
| 29 | <p>Cabarrão, V., Batista, F., Moniz, H., Trancoso, I. & Mata, A. I. (2018). Acoustic-prosodic entrainment in structural metadata events. In Sekhar C.C., Rao P., Ghosh P.K., Murthy H.A., Yegnanarayana B., Umesh S., Alku P., Prasanna S.R.M., Narayanan S. (Ed.), 19th Annual Conference of the International Speech Communication, INTERSPEECH 2018. (pp. 2176-2180). Hyderabad: International Speech Communication Association.</p> <ul style="list-style-type: none"> - Times Cited Web of Science®: 4 - Times Cited Scopus: 3 |
| 30 | <p>Lopes-Teixeira, D., Batista, F. & Ribeiro, R. (2018). Discovering trends in brand interest through topic models . In 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. (pp. 245-252). Sevilha: SciTePress .</p> <ul style="list-style-type: none"> - Times Cited Scopus: 2 - Times Cited Google Scholar: 6 |
| 31 | <p>Barreiro, A. & Batista, F. (2018). Contractions: to align or not to align, that is the question . In First Workshop on Linguistic Resources for Natural Language Processing, Coling 2018. (pp. 122-130). Santa Fe: The Association for Computational Linguistics.</p> <ul style="list-style-type: none"> - Times Cited Google Scholar: 4 |
| 32 | <p>Lopes-Teixeira, D., Batista, F. & Ribeiro, R. (2018). Spatio-temporal analysis of brand interest using social networks. In 13th Iberian Conference on Information Systems and Technologies (CISTI). Cáceres: IEEE.</p> <ul style="list-style-type: none"> - Times Cited Scopus: 3 - Times Cited Google Scholar: 6 |

| | |
|----|---|
| 33 | <p>Soares, C., Batista, F. & Ribeiro, R. (2017). A simplified method to enhance the analysis for new information systems in corporate environments. In 12th Iberian Conference on Information Systems and Technologies, CISTI 2017. (pp. 619-622). Lisbon: IEEE.</p> <p>- Times Cited Google Scholar: 3</p> |
| 34 | <p>Pascoal, R., Ribeiro, R., Batista, F. & de Almeida, A. (2017). Adapting speech recognition in augmented reality for mobile devices in outdoor environments. In Ricardo Queirós and Mário Pinto and Alberto Simões and José Paulo Leal and Maria João Varanda (Ed.), 6th Symposium on Languages, Applications and Technologies (SLATE 2017). Porto: Schloss Dagstuhl- Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing.</p> <p>- Times Cited Scopus: 2</p> <p>- Times Cited Google Scholar: 15</p> |
| 35 | <p>Carvalho, J. P., Rosa, H. & Batista, F. (2017). Detecting relevant tweets in very large tweet collections: the London Riots case study. In 2017 IEEE International Conference on Fuzzy Systems, FUZZ 2017. Naples: IEEE.</p> <p>- Times Cited Scopus: 5</p> <p>- Times Cited Google Scholar: 10</p> |
| 36 | <p>Sousa, M., Trancoso, I., Moniz, H. & Batista, F. (2017). Prosodic exercises for children with ASD via virtual therapy. In Ana Londral, Ana Correia de Barros, Andreia Matos, Célia Sousa, Luís Garcia e Rita Oliveira (Ed.), Atas da Conferência Jornadas SUPERA 2017. (pp. 59-69).: Sociedade Portuguesa de Engenharia de Reabilitação, Tecnologias de Apoio e Acessibilidade.</p> |
| 37 | <p>Solera-Ureña, R., Moniz, H., Batista, F., Cabarrão, R., Pompili, A., Astudillo, R....Trancoso, I. (2017). A semi-supervised learning approach for acoustic-prosodic personality perception in under-resourced domains. In 18th Annual Conference of the International Speech Communication Association, INTERSPEECH 2017. (pp. 929-933).: International Speech Communication Association.</p> <p>- Times Cited Web of Science®: 7</p> <p>- Times Cited Scopus: 7</p> |
| 38 | <p>Jusupova, A., Batista, F. & Ribeiro, R. (2016). Characterizing the personality of Twitter users based on their timeline information. In Azevedo, A., Sarmiento, A., e Rodrigues, L. S. (Ed.), Atas da 16ª Conferência da Associação Portuguesa de Sistemas de Informação - Os Sistemas de Informação na Inovação e na Competitividade. (pp. 292 - 299). Porto: APSI - Associação Portuguesa de Sistemas de Informação, ISCAP/IPP - Instituto Superior de Contabilidade e Administração do Politécnico do Porto.</p> <p>- Times Cited Scopus: 5</p> <p>- Times Cited Google Scholar: 11</p> |
| 39 | <p>Lerner, A., Silber-Varod, V., Batista, F. & Moniz, H. (2016). In search of the role's footprints in client-therapist dialogues. In Barnes, J., Brugos, A., Shattuck-Hufnagel, S., and Veilleux, N. (Ed.), Proceedings of the International Conference on Speech Prosody 2016. (pp. 400-404). Boston, USA: International Speech Communications Association.</p> <p>- Times Cited Google Scholar: 9</p> |
| 40 | <p>Vicente, M., Batista, F. & Carvalho, J. P. (2016). Improving Twitter gender classification using multiple classifiers. In Kóczy, L., and Medina, J. (Ed.), Proceedings of ESCIM 2016. (pp. 121-127). Sofia: Universidad de Cádiz.</p> <p>- Times Cited Google Scholar: 3</p> |
| 41 | <p>Barreiro, A. & Batista, F. (2016). Machine translation of non-contiguous multiword units. In Workshop on Discontinuous Structures in Natural Language Processing (DiscoNLP 2016). (pp. 22-30). San Diego, California, USA: Association for Computational Linguistics.</p> <p>- Times Cited Google Scholar: 12</p> |

| | |
|----|---|
| 42 | <p>Cabarrão, V., Trancoso, I., Mata, A. I., Moniz, H. & Batista, F. (2016). Global analysis of entrainment in dialogues. In Mateo, C. G., Ortega, A., Abad, A., Mamede, N., Martínez Hinarejos, C. D., Teixeira, A., Batista, F., and Perdigão, F. (Ed.), <i>Advances in Speech and Language Technologies for Iberian Languages. IberSPEECH 2016. Lecture Notes in Computer Science.</i> (pp. 215-223). Lisbon: Springer International Publishing.</p> <p>- Times Cited Web of Science®: 9 - Times Cited Google Scholar: 17</p> |
| 43 | <p>Ribeiro, E., Batista, F., Trancoso, I., Lopes, J., Ribeiro, R. & Matos, D. M. de. (2016). Assessing user expertise in spoken dialog system interactions. In Mateo, C. G., Ortega, A., Abad, A., Mamede, N., Martínez Hinarejos, C. D., Teixeira, A., Batista, F., and Perdigão, F. (Ed.), <i>Advances in Speech and Language Technologies for Iberian Languages. IberSPEECH 2016. Lecture Notes in Computer Science.</i> (pp. 245-254). Lisboa: Springer International Publishing.</p> <p>- Times Cited Web of Science®: 1 - Times Cited Scopus: 2 - Times Cited Google Scholar: 5</p> |
| 44 | <p>Ribeiro, E., Batista, F., Trancoso, I., Ribeiro, R. & Matos, D. M. de. (2016). Automatic detection of hyperarticulated speech. In Mateo, C. G., Ortega, A., Abad, A., Mamede, N., Martínez Hinarejos, C. D., Teixeira, A., Batista, F., and Perdigão, F. (Ed.), <i>Advances in Speech and Language Technologies for Iberian Languages. IberSPEECH 2016. Lecture Notes in Computer Science.</i> (pp. 182-191). Lisboa: Springer International Publishing.</p> <p>- Times Cited Web of Science®: 1 - Times Cited Scopus: 1 - Times Cited Google Scholar: 1</p> |
| 45 | <p>Batista, F., Curto, P., Trancoso, I., Abad, A., Ferreira, J., Ribeiro, E....Ribeiro, R. (2016). SPA: Web-based platform for easy access to speech processing modules. In Calzolari, N., Choukri, K., Declerck, T., Goggi, S., Grobelnik, M., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odijk, J., and Piperidis, S. (Ed.), <i>LREC 2016, Tenth International Conference on Language Resources and Evaluation.</i> (pp. 3886-3892). Portorož, Slovenia: ELDA/ELRA.</p> <p>- Times Cited Web of Science®: 1 - Times Cited Scopus: 4 - Times Cited Google Scholar: 5</p> |
| 46 | <p>Solera-Ureña, R., Moniz, H., Batista, F., Astudillo, R. F., Campos, J., Paiva, A....Trancoso, I. (2016). Acoustic-prosodic automatic personality trait assessment for adults and children. In Alberto Abad, Alfonso Ortega and Antonio Teixeira and Carmen Mateo and Carlos Hinarejos and Fernando Perdigão and Fernando Batista and Nuno Mamede (Ed.), <i>Advances in Speech and Language Technologies for Iberian Languages. IberSPEECH 2016. Lecture Notes in Computer Science.</i> (pp. 192-201).: Springer.</p> <p>- Times Cited Web of Science®: 3 - Times Cited Google Scholar: 8</p> |
| 47 | <p>Vicente, M., Batista, F. & Carvalho, J. P. (2016). Creating extended gender labelled datasets of Twitter users. In Carvalho, J. P., Lesot, M.-J., Kaymak, U., Vieira, S., Bouchon-Meunier, B., and Yager, R. R. (Ed.), <i>Information Processing and Management of Uncertainty in Knowledge-Based Systems. Communications in Computer and Information Science.</i> (pp. 690-702). Eindhoven: Springer.</p> <p>- Times Cited Web of Science®: 3 - Times Cited Scopus: 2 - Times Cited Google Scholar: 4</p> |
| 48 | <p>Julião, M., Silva, J., Aguiar, A., Moniz, H. & Batista, F. (2015). Speech features for discriminating stress using branch and bound wrapper search. In José Luis Sierra Rodríguez, José Paulo Leal, Alberto Simões (Ed.), <i>IV Symposium on Languages, Applications and Technologies (SLATE'15), Proceedings.</i> Madrid: Springer.</p> |

| | |
|----|--|
| 49 | <p>Brogueira, G., Batista, F. & Carvalho, J. (2015). Sistema inteligente de recolha, armazenamento e visualização de informação proveniente do twitter. In Atas da 15 Conferência da Associação Portuguesa de Sistemas de Informação (CAPSI 2015). (pp. 418-439). Lisboa: CAPSI.</p> <p>- Times Cited Google Scholar: 7</p> |
| 50 | <p>Ribeiro, E., J. Ferreira, J. Olcoz, Abad, A., Moniz, H., Batista, F....Trancoso, I. (2015). Combining multiple approaches to predict the degree of nativeness. In Interspeech 2015. (pp. 488-492): ISCA.</p> <p>- Times Cited Google Scholar: 11</p> |
| 51 | <p>E. Ribeiro, J. Ferreira, J. Olcoz, Abad, A., Moniz, H., Batista, F....Trancoso, I. (2015). Combining Multiple Approaches to Predict the Degree of Nativeness. In Proc. of Interspeech 2015. Dresden, Germany</p> <p>- Times Cited Web of Science®: 4</p> <p>- Times Cited Scopus: 6</p> <p>- Times Cited Google Scholar: 11</p> |
| 52 | <p>Batista, F. & João P. Carvalho (2015). Text based classification of companies in CrunchBase. In Adnan Yazici, Nikhil R. Pal, Uzat Kaymak (Ed.), IEEE International Fuzzy Systems conference proceedings. Istambul: IEEE.</p> <p>- Times Cited Web of Science®: 5</p> <p>- Times Cited Scopus: 14</p> <p>- Times Cited Google Scholar: 29</p> |
| 53 | <p>Vicente, M., Carvalho, J. P. & Batista, F. (2015). Using unstructured profile information for gender classification of Portuguese and English. In José-Luis Sierra-Rodríguez, José-Paulo Leal, Alberto Simões (Ed.), SLATE 2015: 4th International Symposium on Languages, Applications and Technologies: Languages, Applications and Technologies. (pp. 57-64): Springer.</p> <p>- Times Cited Web of Science®: 5</p> <p>- Times Cited Scopus: 3</p> |
| 54 | <p>J. Lopes, G. Salvi, G. Skantze, Abad, A., J. Gustafson, Batista, F....Trancoso, I. (2015). Detecting Repetitions in Spoken Dialogue Systems Using Phonetic Distances. In Proc. of Interspeech 2015. Dresden, Germany</p> <p>- Times Cited Web of Science®: 8</p> <p>- Times Cited Scopus: 7</p> <p>- Times Cited Google Scholar: 13</p> |
| 55 | <p>Vicente, M., Batista, F. & Carvalho, J. (2015). Twitter gender classification using user unstructured information. In Adnan Yazici, Nikhil R. Pal, Uzat Kaymak (Ed.), 2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE): Proceedings. Istambul: IEEE.</p> <p>- Times Cited Web of Science®: 7</p> <p>- Times Cited Scopus: 22</p> <p>- Times Cited Google Scholar: 49</p> |
| 56 | <p>Gaspar Brogueira, Batista, F. & João P. Carvalho (2015). Arquitetura e Desenvolvimento de um Repositório de Tweets em Português Europeu. In Actas das 5as Jornadas de Informática da Universidade de Évora (JIUE 2015). Évora, Portugal</p> <p>- Times Cited Google Scholar: 4</p> |
| 57 | <p>Gaspar Brogueira, Batista, F. & João P. Carvalho (2015). Using Geolocated Tweets for Characterization of Portuguese Administrative Regions. In In Proc. of 18th AGILE International Conference on Geographic Information Science. Lisbon, Portugal</p> <p>- Times Cited Google Scholar: 3</p> |
| 58 | <p>Moniz, H., A. Pompilli, Batista, F., Trancoso, I., Abad, A. & C. Amorim (2015). Automatic recognition of prosodic patterns in semantic verbal fluency tests - an animal naming task for edutainment applications. In Proc. of ICPHS: 18th International Congress of Phonetic Sciences. Glasgow</p> <p>- Times Cited Google Scholar: 4</p> |

| | |
|----|--|
| 59 | Vera Cabarrão, Moniz, H., J. Ferreira, Batista, F., Trancoso, I. & Mata, Ana I. (2015). Prosodic classification of discourse markers. In Proc. of ICPHS: 18th International Congress of Phonetic Sciences. Glasgow - Times Cited Google Scholar: 14 |
| 60 | Moniz, H., Mata, Ana I., Julia Hirschberg, Batista, F., Andrew Rosenberg & Trancoso, I. (2014). Extending AuToBI to Prominence Detection in European Portuguese. In Speech Prosody. Dublin, Ireland - Times Cited Scopus: 10 - Times Cited Google Scholar: 16 |
| 61 | Barreiro, A., Batista, F., Ribeiro, R., Moniz, H. & Trancoso, I. (2014). OpenLogos Semantico-Syntactic Knowledge-Rich Bilingual Dictionaries. In Nicoletta Calzolari (Conference Chair) and Khalid Choukri and Thierry Declerck and Hrafn Loftsson and Bente Maegaard and Joseph Mariani and Asuncion Moreno and Jan Odijk and Stelios Piperidis (Ed.), Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). (pp. 3774-3781). Reykjavik, Iceland: European Language Resources Association (ELRA). - Times Cited Scopus: 3 - Times Cited Google Scholar: 8 |
| 62 | Gaspar Brogueira, Batista, F., João P. Carvalho & Moniz, H. (2014). Expanding a Database of Portuguese Tweets. In Maria João Varanda Pereira and José Paulo Leal and Alberto Simões (Ed.), 3rd Symposium on Languages, Applications and Technologies. (pp. 275-282). Dagstuhl, Germany: Schloss Dagstuhl--Leibniz-Zentrum fuer Informatik. - Times Cited Scopus: 9 - Times Cited Google Scholar: 13 |
| 63 | Vera Cabarrão, Moniz, H., Batista, F., Ribeiro, R., Mamede, N., Meinedo, H....de Matos, D. (2014). Revising the Annotation of a Broadcast News Corpus: a Linguistic Approach. In Nicoletta Calzolari (Conference Chair) and Khalid Choukri and Thierry Declerck and Hrafn Loftsson and Bente Maegaard and Joseph Mariani and Asuncion Moreno and Jan Odijk and Stelios Piperidis (Ed.), Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). (pp. 3908-3913). Reykjavik, Iceland: European Language Resources Association (ELRA). - Times Cited Web of Science®: 3 - Times Cited Scopus: 2 - Times Cited Google Scholar: 4 |
| 64 | Gaspar Brogueira, Batista, F., João P. Carvalho & Moniz, H. (2014). Portuguese Geolocated Tweets: An Overview. In Proceedings of the International Conference on Information Systems and Design of Communication. (pp. 178-179). New York, NY, USA - Times Cited Web of Science®: 4 - Times Cited Scopus: 3 - Times Cited Google Scholar: 6 |
| 65 | Barreiro, A., Monti, J., Orliac, B., Preuß, S., Arrieta, K., Ling, W....Trancoso, I. (2014). Linguistic evaluation of support verb constructions by OpenLogos and google translate. In Nicoletta Calzolari, Khalid Choukri, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, Stelios Piperidis (Ed.), Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 2014). (pp. 35-40). Reykjavik: European Language Resources Association (ELRA). - Times Cited Web of Science®: 2 - Times Cited Scopus: 4 |

| | |
|----|---|
| 66 | <p>Mata, A., Moniz, H., Mória, T., Gonçalves, A., Silva, F., Batista, F....Falé, I. (2014). Prosodic, syntactic, semantic guidelines for topic structures across domains and corpora. In Nicoletta Calzolari, Khalid Choukri, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, Stelios Piperidis (Ed.), Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). (pp. 1188-1193). Reykjavik: European Language Resources Association (ELRA).</p> <ul style="list-style-type: none"> - Times Cited Web of Science®: 1 - Times Cited Scopus: 2 - Times Cited Google Scholar: 5 |
| 67 | <p>Mata, A., Moniz, H., Batista, F. & Hirschberg, J. (2014). Teenage and adult speech in school context: building and processing a corpus of European Portuguese. In Nicoletta Calzolari, Khalid Choukri, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, Stelios Piperidis (Ed.), Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 2014). (pp. 3914-3919). Reykjavik: European Language Resources Association (ELRA).</p> <ul style="list-style-type: none"> - Times Cited Web of Science®: 3 - Times Cited Google Scholar: 6 |
| 68 | <p>Moniz, H., Batista, F., Mata, Ana I. & Trancoso, I. (2014). Discriminação automática de sinais de pontuação e de disfluências. In Textos Seleccionados do XXIX Encontro Nacional da Associação Portuguesa de Linguística.</p> <ul style="list-style-type: none"> - Times Cited Google Scholar: 1 |
| 69 | <p>Rosa, H., Batista, F. & Carvalho, J. (2014). Twitter topic fuzzy fingerprints. In 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE): Proceedings. Beijing: IEEE.</p> <ul style="list-style-type: none"> - Times Cited Web of Science®: 26 - Times Cited Scopus: 31 - Times Cited Google Scholar: 44 |
| 70 | <p>Hugo Rosa, João P. Carvalho & Batista, F. (2014). Detecting a Tweet's Topic within a Large Number of Portuguese Twitter Trends. In Maria João Varanda Pereira and José Paulo Leal and Alberto Simões (Ed.), 3rd Symposium on Languages, Applications and Technologies. (pp. 185-199). Dagstuhl, Germany: Schloss Dagstuhl--Leibniz-Zentrum fuer Informatik.</p> <ul style="list-style-type: none"> - Times Cited Scopus: 14 - Times Cited Google Scholar: 24 |
| 71 | <p>Barreiro, A., Monti, J., Orliac, B. & Batista, F. (2013). When Multiwords Go Bad in Machine Translation. In Johanna Monti, Ruslan Mitkov, Gloria Corpas Pastor, Violeta Seretan (Ed.), Proceedings of the Workshop on Multi-word Units in Machine Translation and Translation Technology. (pp. 26-33). Nice, France</p> <ul style="list-style-type: none"> - Times Cited Google Scholar: 43 |
| 72 | <p>João P. Carvalho, Calais Pedro, V. & Batista, F. (2013). Towards Intelligent Mining of Public Social Networks' Influence in Society. In IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS). (pp. 478-483). Edmonton, Canada</p> <ul style="list-style-type: none"> - Times Cited Scopus: 8 - Times Cited Google Scholar: 13 |
| 73 | <p>Moniz, H., Batista, F., Trancoso, I. & Mata, Ana I. (2013). Automatic structural metadata identification based on multilayer prosodic information. In Robert Eklund (Ed.), Proceedings of Disfluency in Spontaneous Speech, DiSS 2013. (pp. 49-52). Stockholm, Sweden</p> <ul style="list-style-type: none"> - Times Cited Scopus: 1 - Times Cited Google Scholar: 4 |

| | |
|----|---|
| 74 | <p>Medeiros, H., Batista, F., Moniz, H., Trancoso, I. & Nunes, L. (2013). Comparing different machine learning approaches for disfluency structure detection in a corpus of university lectures. In Leal, J. P., Rocha, R., and Simões, A. (Ed.), 2nd Symposium on Languages, Applications and Technologies, SLATE 2013. (pp. 259-269). Porto: OASlcs.</p> <ul style="list-style-type: none"> - Times Cited Scopus: 1 - Times Cited Google Scholar: 5 |
| 75 | <p>Medeiros, H., Moniz, H., Batista, F., Trancoso, I. & Nunes, L. (2013). Disfluency detection based on prosodic features for university lectures. In Bimbot, F., Cerisara, C., Fougeron, C., Gravier, G., Lamel, L., Pellegrino, F., and Perrier, P. (Ed.), Proceedings of the 14th Annual Conference of the International Speech Communication Association (INTERSPEECH 2013) . (pp. 2629-2633). Lyon: International Speech Communication Association.</p> <ul style="list-style-type: none"> - Times Cited Web of Science®: 8 - Times Cited Scopus: 17 - Times Cited Google Scholar: 24 |
| 76 | <p>João P. Carvalho, Batista, F. & L. Coheur (2012). A Critical Survey on the use of Fuzzy Sets in Speech and Natural Language Processing. In IEEE World Congress on Computational Intelligence (IEEE-WCCI). (pp. 270-277). Brisbane, Australia</p> <ul style="list-style-type: none"> - Times Cited Web of Science®: 4 - Times Cited Scopus: 19 - Times Cited Google Scholar: 37 |
| 77 | <p>Moniz, H., Batista, F., Trancoso, I. & Mata, Ana I. (2012). Análise de interrogativas em diferentes domínios. In Textos Seleccionados do XXVII Encontro Nacional da APL. (pp. 0-0). Lisboa</p> <ul style="list-style-type: none"> - Times Cited Google Scholar: 4 |
| 78 | <p>Moniz, H., Batista, F., Mata, Ana I. & Trancoso, I. (2012). Analysis of disfluencies in a corpus of university lectures. In 5th ISEL Conf. on Experimental Linguistics. (pp. 89-92). Athens</p> <ul style="list-style-type: none"> - Times Cited Google Scholar: 8 |
| 79 | <p>Pellegrini, T, Moniz, H., Batista, F., Trancoso, I. & Astudillo, R. (2012). Extension of the LECTRA corpus: classroom LECTure TRANscriptions in European Portuguese. In Heliana Mello, Massimo Pettorino, Tommaso Raso (Ed.), Speech and Corpora. (pp. 98-102). Belo Horizonte: Firenze University Press.</p> <ul style="list-style-type: none"> - Times Cited Google Scholar: 10 |
| 80 | <p>Moniz, H., Batista, F., Meinedo, H., Abad, A., Trancoso, I., Mata, Ana I....Mamede, N. (2010). Prosodically-based automatic segmentation and punctuation. In Speech prosody 2010 (Ed.), Speech Prosody 2010. (pp. 1-4): Speech prosody 2010.</p> <ul style="list-style-type: none"> - Times Cited Google Scholar: 17 |
| 81 | <p>Batista, F., Moniz, H., Trancoso, I., Meinedo, H., Mata, Ana I. & Mamede, N. (2010). Extending the punctuation module for European Portuguese. In ISCA (Ed.), Interspeech 2010. (pp. 1509-1512). Makuhari: ISCA.</p> <ul style="list-style-type: none"> - Times Cited Web of Science®: 3 - Times Cited Scopus: 8 - Times Cited Google Scholar: 16 |
| 82 | <p>Batista, F., Trancoso, I. & Mamede, N. (2009). Comparing automatic rich transcription for portuguese, spanish and english broadcast news. In Automatic Speech Recognition \& Understanding, 2009. ASRU 2009. IEEE Workshop on. (pp. 540-545).</p> <ul style="list-style-type: none"> - Times Cited Web of Science®: 3 - Times Cited Scopus: 5 - Times Cited Google Scholar: 10 |

| | |
|----|--|
| 83 | Batista, F., Trancoso, I. & Mamede, N. (2009). Automatic Recovery of Punctuation Marks and Capitalization Information for Iberian Languages. In I Joint SIG-IL/Microsoft Workshop on Speech An Language Technologies for Iberian Languages, Porto Salvo, Portugal. (pp. 99-102). - Times Cited Google Scholar: 7 |
| 84 | Batista, F., Mamede, N. & Trancoso, I. (2008). The impact of Language Dynamics on the Capitalization of Broadcast News. In Interspeech 2008. - Times Cited Scopus: 4 - Times Cited Google Scholar: 11 |
| 85 | Mendes, Ana Cristina, L. Coheur, Mamede, N., Ribeiro, R., de Matos, D. & Batista, F. (2008). QA@L2F, first steps at QA@CLEF. In Advances in Multilingual and Multimodal Information Retrieval. (pp. 356-363): Springer-Verlag. - Times Cited Scopus: 6 - Times Cited Google Scholar: 14 |
| 86 | Batista, F., Mamede, N. & Trancoso, I. (2008). Language Dynamics and Capitalization using Maximum Entropy. In Proc. of ACL-08: HLT, Short Papers. (pp. 1-4): ACL. - Times Cited Google Scholar: 14 |
| 87 | Batista, F., R. Amaral, Trancoso, I. & Mamede, N. (2008). Impact of dynamic model adaptation beyond speech recognition. In Proc. of SLT 2008. - Times Cited Google Scholar: 3 |
| 88 | Batista, F., Mamede, N. & Trancoso, I. (2008). Temporal Issues and Recognition Errors on the Capitalization of Speech Transcriptions. In 11th International Conference on Text, Speech and Dialogue (TSD 2008). (pp. 45-52): Springer-Verlag. - Times Cited Google Scholar: 1 |
| 89 | Batista, F., D. Caseiro, Mamede, N. & Trancoso, I. (2007). Recovering Punctuation Marks for Automatic Speech Recognition. In Interspeech 2007. (pp. 1977 - 1980). Antwerp, Belgium - Times Cited Web of Science®: 2 - Times Cited Scopus: 6 - Times Cited Google Scholar: 34 |
| 90 | Mendes, Ana Cristina, L. Coheur, Mamede, N., L. Romão, J. Loureiro, Ribeiro, R....de Matos, D. (2007). QA@L2F@QA@CLEF. In Working Notes for the CLEF 2007 Workshop. |
| 91 | J. Baptista, Batista, F. & Mamede, N. (2006). Building a dictionary of anthroponyms. In Computational Processing of the Portuguese Language. (pp. 21-30): Springer Berlin Heidelberg. - Times Cited Web of Science®: 1 - Times Cited Scopus: 3 - Times Cited Google Scholar: 14 |
| 92 | Ribeiro, R., Batista, F., Paulo, J., Mamede, N. & Pinto, H. (2006). Cooking an ontology. In Euzenat, Jérôme and Domingue, John (Ed.), AIMSA 2006: Artificial intelligence: methodology, systems, and applications, Conference proceedings. (pp. 213-221): Springer Berlin Heidelberg. - Times Cited Web of Science®: 17 - Times Cited Scopus: 36 - Times Cited Google Scholar: 66 |
| 93 | L. Coheur, Batista, F. & Mamede, N. (2004). Towards a flexible syntax/semantics interface. In Proceedings of the Herramientas y Recursos Linguísticos para el Espanol y el Portugués workshop. (pp. 256-272). - Times Cited Google Scholar: 1 |

| | |
|----|---|
| 94 | L. Coheur, Batista, F. & J. Paulo (2003). JaVaLI!: understanding real questions. In EUROLAN 2003 student workshop: Applied Natural Language Processing, possible applications for the Semantic Web, Bucharest, Romania, July. - Times Cited Google Scholar: 4 |
| 95 | Batista, F. & Mamede, N. (2002). SuSAna: Módulo multifuncional da análise sintáctica de superf{v}cie. In Proc. Multilingual Information Access and Natural Language Processing Workshop (IBERAMIA 2002). (pp. 29-37). |
| 96 | L. Wittmann, Ribeiro, R., T. Pêgo & Batista, F. (2000). Some Language Resources and Tools for Computational Processing of Portuguese at INESC. In Proceedings of the 2nd International Conference on Language Resources and Evaluation - LREC 2000.: ELRA. - Times Cited Scopus: 2 - Times Cited Google Scholar: 11 |

- Conference proceedings editor

| | |
|---|--|
| 1 | M.J. Lesot, S. Vieira, Marek Z. Reformat, João P. Carvalho, Batista, F., Bernadette Bouchon-Meunier...Ronald R. Yager (2025). Information Processing and Management of Uncertainty in Knowledge-Based Systems, Volume 3. Springer Science and Business Media LLC. |
| 2 | M.J. Lesot, S. Vieira, M. Reformat, Batista, F., João P. Carvalho, Bernadette Bouchon-Meunier...Ronald R. Yager (2025). IPMU2024 Lisboa - Short Paper Proceedings. Lisboa. Zenodo. |
| 3 | M.J. Lesot, S. Vieira, M. Reformat, João P. Carvalho, Batista, F., Bernadette Bouchon-Meunier...Ronald R. Yager (2024). Information Processing and Management of Uncertainty in Knowledge-Based Systems: 20th International Conference, IPMU 2024, Lisbon, Portugal, July 22-26, 2024, Proceedings, Volume 1. Lisbon. Springer Science and Business Media LLC. |
| 4 | Vlória Pinheiro, Pablo Gamallo, Raquel Amaro, Carolina Scarton, Batista, F., Diego Silva...Hugo Pinto (2022). Computational Processing of the Portuguese Language: 15th International Conference, PROPOR 2022, Fortaleza, Brazil, March 21–23, 2022, Proceedings. Fortaleza, Brazil. Springer International Publishing. - Times Cited Google Scholar: 3 |
| 5 | Paulo Quaresma, Renata Vieira, Sandra Aluísio, Moniz, H., Batista, F. & Teresa Gonçalves (2020). Computational Processing of the Portuguese Language: 14th International Conference, PROPOR 2020, Evora, Portugal, March 24, 2020, Proceedings. Évora. Springer International Publishing. |
| 6 | A. Martins, Moniz, H., S. Fumega, Bruno Martins, Batista, F., L. Coheur...M. Forcada (2020). Proceedings of the 22nd Annual Conference of the European Association for Machine Translation. Online (in place of Lisbon). European Association for Machine Translation. - Times Cited Google Scholar: 4 |
| 7 | Ricardo Rodrigues, Jan Janoušek, Luís Ferreira, L. Coheur, Batista, F. & Hugo Oliveira (2019). 8th Symposium on Languages, Applications and Technologies. Schloss Dagstuhl--Leibniz-Zentrum fuer Informatik. |
| 8 | Abad, A., A. Ortega, António Teixeira, C. Mateo, C. Hinarejos, F. Perdigão...Mamede, N. (2016). Advances in Speech and Language Technologies for Iberian Languages. Springer International Publishing. - Times Cited Scopus: 1 |

- Talk

| | |
|----|---|
| 1 | Medina, Roberto Carlos, Batista, F. & Lima, Estanislau (2025). CrioletSet: A New Corpus for Cape Verdean Creole, Towards Robust Machine Translation. Progress in Artificial Intelligence: 24th EPIA Conference on Artificial Intelligence, EPIA 2025. |
| 2 | Cardoso, M., Ribeiro, E. & Batista, F. (2025). Portuguese Far-Right Discourse on Social Media: Insights from Topic Modeling. 14th Symposium on Languages, Applications and Technologies (SLATE 2025). |
| 3 | Ramos, M. & Batista, F. (2024). A utilização de Plataformas de IA por estudantes de licenciatura. CIDU 2024 - XIII Congresso Ibero-americano de Docência Universitária. |
| 4 | Ramos, M. & Batista, F. (2024). Plágio, IA e conduta académica: conceitos, práticas e novos desafios. Sessões Pedagógicas, Iscte. |
| 5 | Fonseca, A., Moro, S., pontes, C., Batista, F., Ribeiro, R., Marques, C....Guerra, R. (2024). Characterization of user positioning in online social media, application to hate speech conversations. NetSci - International School and Conference on Network Science. |
| 6 | Moro, S., Fonseca, A., pontes, C., Batista, F., Ribeiro, R., Marques, C....Guerra, R. (2024). Unveiling research on social networks in the Portuguese language. NetSci - International School and Conference on Network Science. |
| 7 | Espírito-Santo, Ana, Santos & Batista, F. (2024). Unveiling Online Violence Against Women in Politics: EU Political Candidates. IPSA 75th Anniversary Conference. |
| 8 | Espírito-Santo, Ana, Santos & Batista, F. (2024). Unveiling Online Violence Against Women in Politics: EU Political Candidates. 120th APSA Annual Meeting and Exhibition. |
| 9 | Marques, C., Guerra, R., Carvalho, P., Sarroeira, R., Fonseca, A., Batista, F....Moro, S. (2024). Online Hate Speech: How discursive phenomena and rhetorical mechanisms influence negative emotions?. XXXI Meeting of the Portuguese Association of Classification and Data Analysis (JOCLAD 2024) . |
| 10 | Rodrigues, D., Lopes, A. L. & Batista, F. (2023). Web of Science Citation Gaps: An Automatic Approach to Detect Indexed but Missing Citations. Symposium on Languages, Applications and Technologies. |
| 11 | Ricardo Rei, Batista, F., Nuno Miguel Guerreiro & L. Coheur (2021). Multilingual Simultaneous Sentence End and Punctuation Prediction. Swiss Text Analytics Conference 2021. |
| 12 | Afonso Pinto, Moniz, H. & Batista, F. (2020). Detection of Emerging Words in Portuguese Tweets. 9th Symposium on Languages, Applications and Technologies (SLATE 2020). |
| 13 | Moniz, H., R. Solera-Ureña, Vera Cabarrão, M. Julião, Batista, F. & Trancoso, I. (2019). Affective computing based on acoustic-prosodic cues. --. |
| 14 | Vera Cabarrão, M. Julião, R. Solera-Ureña, Moniz, H., Batista, F., Trancoso, I....Mata, Ana I. (2019). Affective analysis of customer service calls. 10th International Conference of Experimental Linguistics (ExLing 2019). |
| 15 | Diana Lopes-Teixeira, Batista, F. & Ribeiro, R. (2018). Discovering trends in brand interest through topic models. 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. |
| 16 | Carlos Soares, Batista, F. & Ribeiro, R. (2017). A Simplified Method to Enhance the Analysis for new Information Systems in Corporate Environments. Proc. of CISTI'2017: 12th Iberian Conference on Information Systems and Technologies. |

| | |
|----|--|
| 17 | Rui Pascoal, Ribeiro, R., Batista, F. & de Almeida, A. (2017). Adapting Speech Recognition in Augmented Reality for Mobile Devices in Outdoor Environments. 6th Symposium on Languages, Applications and Technologies (SLATE 2017). |
| 18 | Batista, F., Pedro Curto, Trancoso, I., Abad, A., J. Ferreira, E. Ribeiro...Ribeiro, R. (2016). SPA: web-based platform for easy access to speech processing modules. 10th Language Resources and Evaluation Conference - LREC 2016. |
| 19 | Barreiro, A. & Batista, F. (2016). Machine Translation of Non-Contiguous Multiword Units. DiscoNLP 2016 - NAACL HLT 2016 Workshop on Discontinuous Structures in NLP. 22-30 |
| 20 | Anat Lerner, Vered Silber-Varod, Batista, F. & Moniz, H. (2016). In search of the role's footprints in client-therapist dialogues. Proceedings of the International Conference on Speech Prosody. 400-404 - Times Cited Scopus: 4 |
| 21 | E. Ribeiro, Batista, F., Trancoso, I., J. Lopes, Ribeiro, R. & de Matos, D. (2016). Assessing User Expertise in Spoken Dialog System Interactions. IberSPEECH 2016. |
| 22 | R. Solera-Ureña, Moniz, H., Batista, F., R. Astudillo, J. Campos, Paiva, A....Trancoso, I. (2016). Acoustic-Prosodic Automatic Personality Trait Assessment for Adults and Children. IberSPEECH 2016. - Times Cited Web of Science®: 1 - Times Cited Scopus: 1 |
| 23 | Batista, F., E. Ribeiro, Trancoso, I., Ribeiro, R. & de Matos, D. (2016). Automatic Detection of Hyperarticulated Speech. IberSPEECH 2016. |
| 24 | Vera Cabarrão, Trancoso, I., Mata, Ana I., Batista, F. & Moniz, H. (2016). Global analysis of entrainment in dialogues. IberSPEECH 2016, Lisbon, Portugal, November 23-25, 2016. - Times Cited Scopus: 7 |
| 25 | F. Rebelo, Batista, F. & Ribeiro, R. (2016). Cascatas de Classificação de Sentimento em Microblogs. INForum 2016, Gestão de Dados e Conhecimento. 203-214 |
| 26 | M. Vicente, Batista, F. & João P. Carvalho (2016). Improving Twitter gender classification using multiple classifiers. Proc. of the 8th European Symposium on Computational Intelligence and Mathematics (ESCIM 2016). 121-127 |
| 27 | A. Jusupova, Batista, F. & Ribeiro, R. (2016). Characterizing the Personality of Twitter Users based on their Timeline Information. Atas da 16 Conferência da Associação Portuguesa de Sistemas de Informação. 292 - 299 - Times Cited Google Scholar: 3 |
| 28 | M. Vicente, Batista, F. & João P. Carvalho (2016). Creating Extended Gender Labelled Datasets of Twitter Users. IPMU2016 - 16th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based systems. |
| 29 | Dias, L., Brandão, T. & Batista, F. (2016). Detecting violence on movie excerpts - A machine-learning approach based on audio and video features. INForum 2016. --, ----- |
| 30 | E. Ribeiro, J. Ferreira, J. Olcoz, Abad, A., Moniz, H., Batista, F....Trancoso, I. (2015). Combining Multiple Approaches to Predict the Degree of Nateness. Proc. of Interspeech 2015. |

| | |
|----|---|
| 31 | M. Julião, J. Silva, A. Aguiar, Moniz, H. & Batista, F. (2015). Speech Features for Discriminating Stress Using Branch and Bound Wrapper Search. IV Symposium on Languages, Applications and Technologies (SLATE'15). 1-9 |
| 32 | J. Lopes, G. Salvi, G. Skantze, Abad, A., J. Gustafson, Batista, F....Trancoso, I. (2015). Detecting Repetitions in Spoken Dialogue Systems Using Phonetic Distances. Proc. of Interspeech 2015. |
| 33 | M. Vicente, Batista, F. & João P. Carvalho (2015). Twitter gender classification using user unstructured information. Proc. of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). |
| 34 | Batista, F. & João P. Carvalho (2015). Text based classification of companies in CrunchBase. IEEE International Conference on Fuzzy Systems (FUZZ-IEEE2015). - Times Cited Web of Science®: 2 |
| 35 | Gaspar Brogueira, Batista, F. & João P. Carvalho (2015). Arquitetura e Desenvolvimento de um Repositório de Tweets em Português Europeu. Actas das 5as Jornadas de Informática da Universidade de Évora (JIUE 2015). |
| 36 | Gaspar Brogueira, Batista, F. & João P. Carvalho (2015). Using Geolocated Tweets for Characterization of Portuguese Administrative Regions. In Proc. of 18th AGILE International Conference on Geographic Information Science. |
| 37 | M. Vicente, João P. Carvalho & Batista, F. (2015). Using Unstructured Profile Information for Gender Classification of Portuguese and English Twitter users. IV Symposium on Languages, Applications and Technologies (SLATE'15). 1-6 |
| 38 | M. Vicente, João P. Carvalho & Batista, F. (2015). Using Unstructured Profile Information for Gender Classification of English and Portuguese Twitter users. Symposium on Languages, Applications and Technologies (SLATE'15). |
| 39 | Moniz, H., A. Pompilli, Batista, F., Trancoso, I., Abad, A. & C. Amorim (2015). Automatic recognition of prosodic patterns in semantic verbal fluency tests - an animal naming task for edutainment applications. Proc. of ICPHS: 18th International Congress of Phonetic Sciences. |
| 40 | M. Julião, J. Silva, A. Aguiar, Moniz, H. & Batista, F. (2015). Speech Features for Discriminating Stress Using Bound and Branch Wrapper Search. Symposium on Languages, Applications and Technologies (SLATE'15). |
| 41 | Gaspar Brogueira, Batista, F. & João P. Carvalho (2015). Sistema Inteligente de Recolha, Armazenamento e Visualizacao de Informacao proveniente do Twitter. Atas da 15 Conferência da Associação Portuguesa de Sistemas de Informação. 535-556 |
| 42 | Vera Cabarrão, Moniz, H., J. Ferreira, Batista, F., Trancoso, I. & Mata, Ana I. (2015). Prosodic classification of discourse markers. Proc. of ICPHS: 18th International Congress of Phonetic Sciences. |
| 43 | Gaspar Brogueira, Batista, F., João P. Carvalho & Moniz, H. (2014). Expanding a Database of Portuguese Tweets. 3rd Symposium on Languages, Applications and Technologies. 38, 275-282 |
| 44 | Gaspar Brogueira, Batista, F., João P. Carvalho & Moniz, H. (2014). Portuguese Geolocated Tweets: An Overview. Proceedings of the International Conference on Information Systems and Design of Communication. - Times Cited Web of Science®: 3 |

| | |
|----|--|
| 45 | Vera Cabarrão, Moniz, H., Batista, F., Ribeiro, R., Mamede, N., Meinedo, H....de Matos, D. (2014). Revising the Annotation of a Broadcast News Corpus: a Linguistic Approach. Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). 3908-3913 |
| 46 | Hugo Rosa, Batista, F. & João P. Carvalho (2014). Twitter Topic Fuzzy Fingerprints. WCCI2014, FUZZ-IEEE, 2014 IEEE World Congress on Computational Intelligence, International Conference on Fuzzy Systems. 776-783 |
| 47 | Mata, Ana I., Moniz, H., Batista, F. & Julia Hirschberg (2014). Teenage and Adult Speech in School Context: Building and Processing a Corpus of European Portuguese. Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). 3914-3919 - Times Cited Web of Science®: 1 - Times Cited Scopus: 2 |
| 48 | Mata, Ana I., Moniz, H., Telmo Mória, Anabela Gonçalves, Fátima Silva, Batista, F....Isabel Falé (2014). Prosodic, Syntactic, Semantic Guidelines for Topic Structures Across Domains and Corpora. Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). 1188-1193 - Times Cited Web of Science®: 1 |
| 49 | Barreiro, A., Monti, J., Orliac, B., Susanne Preuß, Kutz Arrieta, Wang Ling...Trancoso, I. (2014). Linguistic Evaluation of Support Verb Constructions by OpenLogos and Google Translate. Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). 35-40 - Times Cited Google Scholar: 20 |
| 50 | Luis Silva, Ribeiro, R. & Batista, F. (2014). Portugal by regions: The contrast between social networks and statistical data reference. INFORUM 2014. |
| 51 | Vera Cabarrão, Moniz, H., Batista, F., Trancoso, I., Mata, Ana I. & Curto, S. (2014). Discourse markers in spontaneous speech in European Portuguese: a first approach. International Workshop - Pragmatic Markers, Discourse Markers and Modal Particles. |
| 52 | Moniz, H., Batista, F., Trancoso, I. & Mata, Ana I. (2014). Analytics of conversational fillers in European Portuguese. International Workshop - Pragmatic Markers, Discourse Markers and Modal Particles. |
| 53 | Mata, Ana I., Moniz, H., Telmo Mória, Anabela Gonçalves, Fátima Silva, Batista, F....Isabel Falé (2014). A relação prosódica-sintaxe-discurso: anotação e análise de estruturas de tópico em corpora de fala espontânea. XXX ENCONTRO NACIONAL DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA. |
| 54 | Moniz, H., Mata, Ana I., Julia Hirschberg, Batista, F., Andrew Rosenberg & Trancoso, I. (2014). Predição de proeminência prosódica com base no sistema automático AuToBI. XXX ENCONTRO NACIONAL DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA. |
| 55 | Mata, Ana I., Moniz, H. & Batista, F. (2014). Comparing phrase-final patterns across speech styles and age groups in European Portuguese. 3rd SWIP - Swiss Workshop on Prosody. |
| 56 | Hugo Rosa, João P. Carvalho & Batista, F. (2014). Detecting a Tweet's Topic within a Large Number of Portuguese Twitter Trends. 3rd Symposium on Languages, Applications and Technologies. 38, 185-199 |
| 57 | Moniz, H., Mata, Ana I., Julia Hirschberg, Batista, F., Andrew Rosenberg & Trancoso, I. (2014). Extending AuToBI to Prominence Detection in European Portuguese. Speech Prosody. |
| 58 | Barreiro, A., Batista, F., Ribeiro, R., Moniz, H. & Trancoso, I. (2014). OpenLogos Semantico-Syntactic Knowledge-Rich Bilingual Dictionaries. Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14). 3774-3781 |

| | |
|----|--|
| 59 | Moniz, H. & Batista, F. (2013). Towards automatic speech processing and intonational labeling for different speaking styles. PaPI 2013. |
| 60 | Barreiro, A., Monti, J., Orliac, B. & Batista, F. (2013). When Multiwords Go Bad in Machine Translation. Proceedings of the Workshop on Multi-word Units in Machine Translation and Translation Technology. 26-33 |
| 61 | João P. Carvalho, Calais Pedro, V. & Batista, F. (2013). Towards Intelligent Mining of Public Social Networks' Influence in Society. IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS). 478-483 |
| 62 | Henrique Medeiros, Moniz, H., Batista, F., Trancoso, I. & Meinedo, H. (2013). Experiments on automatic detection of filled pauses using prosodic features. Proceedings of INForum 2013. Track: Gestão de Dados e Conhecimento. 335 - 345 |
| 63 | Moniz, H., Batista, F., Trancoso, I. & Mata, Ana I. (2013). Automatic structural metadata identification based on multilayer prosodic information. Proceedings of Disfluency in Spontaneous Speech, DiSS 2013. 54, 49-52 |
| 64 | Batista, F. & Ribeiro, R. (2013). Sentiment Analysis and Topic Classification: Case study over Spanish tweets. NESS Conference on Media Opinion Dynamics. - Times Cited Google Scholar: 2 |
| 65 | Ribeiro, R. & Batista, F. (2013). Natural Language Processing for the Social Sciences. NESS Conference on Media Opinion Dynamics. |
| 66 | Moniz, H., Batista, F., Trancoso, I. & Mata, Ana I. (2012). Prosodic context-based analysis of disfluencies. Interspeech 2012. 0-0 |
| 67 | Batista, F. & Mamede, N. (2012). Recovering Capitalization and Punctuation Marks on Speech Transcriptions. PROPOR 2012 - Thesis contest. |
| 68 | João P. Carvalho, Batista, F. & L. Coheur (2012). A Critical Survey on the use of Fuzzy Sets in Speech and Natural Language Processing. IEEE World Congress on Computational Intelligence (IEEE-WCCI). 270-277 |
| 69 | Batista, F. & Ribeiro, R. (2012). The L2F Strategy for Sentiment Analysis and Topic Classification. Workshop on Sentiment Analysis at SEPLN. - Times Cited Google Scholar: 7 |
| 70 | Moniz, H., Batista, F., Mata, Ana I. & Trancoso, I. (2012). Analysis of disfluencies in a corpus of university lectures. ExLing 2012. |
| 71 | Batista, F. (2011). Recovering Capitalization and Punctuation Marks on Speech Transcriptions. PhD Thesis defense. |
| 72 | Moniz, H., Batista, F., Trancoso, I. & Mata, Ana I. (2011). Análise de interrogativas em diferentes domínios. XXVII Encontro Nacional da Associação Portuguesa de Linguística. |
| 73 | Batista, F., Trancoso, I. & Mamede, N. (2009). Comparing automatic rich transcription for portuguese, spanish and english broadcast news. Automatic Speech Recognition & Understanding, 2009. ASRU 2009. IEEE Workshop on. 540-545 |
| 74 | Batista, F., Trancoso, I. & Mamede, N. (2009). Automatic Recovery of Punctuation Marks and Capitalization Information for Iberian Languages. I Joint SIG-IL/Microsoft Workshop on Speech An Language Technologies for Iberian Languages, Porto Salvo, Portugal. 99-102 |

| | |
|----|--|
| 75 | Batista, F., Mamede, N. & Trancoso, I. (2008). Language Dynamics and Capitalization using Maximum Entropy. Proc. of ACL-08: HLT, Short Papers. 1-4 |
| 76 | Batista, F., Mamede, N. & Trancoso, I. (2008). Temporal Issues and Recognition Errors on the Capitalization of Speech Transcriptions. 11th International Conference on Text, Speech and Dialogue (TSD 2008). 5246, 45-52 |
| 77 | Batista, F., Mamede, N. & Trancoso, I. (2008). The impact of Language Dynamics on the Capitalization of Broadcast News. Interspeech 2008. |
| 78 | Batista, F., D. Caseiro, Mamede, N. & Trancoso, I. (2007). Recovering Punctuation Marks for Automatic Speech Recognition. Interspeech 2007. 1977 - 1980 |
| 79 | Batista, F., Mamede, N., D. Caseiro & Trancoso, I. (2007). A Lightweight on-the-fly Capitalization System for Automatic Speech Recognition. Proc. RANLP'07. |
| 80 | Ribeiro, R., Batista, F., J. Paulo, Mamede, N. & Pinto, H. Sofia (2006). Cooking an Ontology. Artificial Intelligence: Methodology, Systems, and Applications. 4183, 213-221 |
| 81 | J. Baptista, Batista, F. & Mamede, N. (2006). Building a dictionary of anthroponyms. Computational Processing of the Portuguese Language. 21-30 |
| 82 | L. Coheur, Batista, F. & Mamede, N. (2004). Towards a flexible syntax/semantics interface. Proceedings of the Herramientas y Recursos Linguísticos para el Espanol y el Portugués workshop. 256-272 |
| 83 | L. Coheur, Batista, F. & J. Paulo (2003). JaValII: understanding real questions. EUROLAN 2003 student workshop: Applied Natural Language Processing, possible applications for the Semantic Web, Bucharest, Romania, July. |
| 84 | Batista, F. & Mamede, N. (2002). SuSAna: Módulo multifuncional da análise sintáctica de superfícies. Proc. Multilingual Information Access and Natural Language Processing Workshop (IBERAMIA 2002). 29-37 |

• Other Publications

- Working Papers

| | |
|---|---|
| 1 | Batista, F. & Mamede, N. (2012). Recovering Capitalization and Punctuation Marks on Speech Transcriptions. PROPOR 2012 web site. 0-0 - Times Cited Google Scholar: 21 |
| 2 | Batista, F., Paulo, J., Mamede, N., Ribeiro, R. & Vaz, Paula Cristina (2006). Ontology construction: cooking domain. Ontology construction: cooking domain. - Times Cited Google Scholar: 47 |

- Other publications

| | |
|---|---|
| 1 | Taborda, B., de Almeida, A., Dias, J. C., Batista, F. & Ribeiro, R. (2021). Stock Market Tweets Data. IEEE Dataport. |
| 2 | João P. Carvalho, M. Reformat, M.J. Lesot, S. Vieira & Batista, F. (2020). Book of Abstracts of IPMU 2020. IPMU 2020. |

Research Projects

| Project Title | Role in Project | Partners | Period |
|---|-------------------|---|-------------|
| Your Own Developmental Agent: Agent-led feedback for improving team interpersonal processes | Researcher | BRU-Iscte (OB&HR), INESC-ID - (Portugal), Critical - Collaborative Institution (Portugal) | 2025 - 2028 |
| European Observatory of Smart Initiatives and Tools for Tourism | Researcher | ISTAR-Iscte - Leader, INESC-ID - (Portugal), NEST - Consultant (Portugal) | 2025 - 2026 |
| Espaço para crescer | Researcher | CIES-Iscte | 2023 - 2024 |
| Artificial Intelligence and Data Science for Public Administration Portugal Innovation Hub | Researcher | ISTAR-Iscte (DLS) - Leader, AMA - (Portugal), AIP - CCI - (Portugal), ANPME - (Portugal), AUDAX - (Portugal), CMS - (Portugal), CMV - (Portugal), CCDR Algarve - (Portugal), CoLABOR - - (Portugal), DGEEC - (Portugal), Esri Portugal - (Portugal), GEP/MTSSS - (Portugal), INDEG - (Portugal), IPPS - (Portugal), Mentortec - (Portugal), MORE CoLAB - (Portugal), NOVA IMS - (Portugal), Oeste CIM - (Portugal), UNINOVA - - (Portugal), UNU-EGOV - (Portugal) | 2023 - 2025 |
| kNOwing online HATE speech: knowledge + awareness = TacklingHate | Researcher | CIS-Iscte (PsyChange) - Leader, BRU-Iscte, ISTAR-Iscte, CIES-Iscte, INESC-ID - (Portugal), IST-ID - (Portugal), ILGA Portugal - (Portugal), CBL - (Portugal), CICDR/ACM - (Portugal), SOS RACISMO - (Portugal) | 2022 - 2024 |
| AppRecommender: Intelligent App Distribution towards an Optimised App Discover | Local Coordinator | Aptoide - Leader (Portugal), Caixa Mágica - Leader (Portugal) | 2019 - 2021 |

Academic Management Positions

Director (2025 - 2027)
Unit/Area: Bachelor Degree in Data Science (PL)

Director (2025 - 2027)
Unit/Area: Bachelor Degree in Data Science (PL)

Director (2025 - 2027)
Unit/Area: Bachelor Degree in Data Science

Director (2023 - 2025)
Unit/Area: Bachelor Degree in Data Science (PL)

Director (2023 - 2025)
Unit/Area: Bachelor Degree in Data Science

Coordenador do 3º Ano (2022 - 2023)
Unit/Area: Bachelor Degree in Data Science

| |
|---|
| Coordenador do 3º Ano (2021 - 2023) Unit/Area: Bachelor Degree in Data Science (PL) |
| Presidente (2017 - 2019) Unit/Area: Conselho Pedagógico |
| Presidente (2017 - 2019) Unit/Area: Comissão Permanente do Conselho Pedagógico |
| Presidente (2017 - 2019) Unit/Area: Plenário do Conselho Pedagógico |
| Membro (Docente) (2015 - 2017) Unit/Area: Comissão Pedagógica |
| Membro (Docente) (2015 - 2017) Unit/Area: Comissão Permanente do Conselho Pedagógico |
| Membro (Docente) (2015 - 2017) Unit/Area: Plenário do Conselho Pedagógico |

Awards

| |
|--|
| One of the winners of the Shared Task "Sentence End and Punctuation Prediction" - Rei et al. (2021) Multilingual Simultaneous Sentence End and Punctuation Prediction. Swiss Text Analytics Conference 2021 (2021) |
| Best Poster Award - Bunga et al. (2021) From Implicit Preferences to Ratings: Video Games Recommendation based on Collaborative Filtering. 13th International Joint Conference on Knowledge Discovery and Information Retrieval (IC3K 2021) (2021) |
| PhD Dissertation - Honorable Mention - Fernando Batista and Nuno Mamede (2012) Recovering Capitalization and Punctuation Marks on Speech Transcriptions. PhD Dissertation contest, PROPOR 2012 (2012) |
| Best Regular Paper Award - Batista et al. (2009) Automatic Recovery of Punctuation Marks and Capitalization Information for Iberian Languages. I Joint SIG-IL/Microsoft Workshop on Speech An Language Technologies for Iberian Languages (2009) |

Professional Associations

| |
|---------------------------------|
| IEEE Senior member (Since 2016) |
| ISCA Speech (Since 2008) |

Organization/Coordination of Events

| Type of Organization/Coordination | Event Title | Organizer | Year |
|--|--|---|------|
| Member of free course/summer school's organizing committee | LxMLS 2026 - The 16th Lisbon Machine Learning School | IST; IT; INESC-ID; LUMLIS, Unbabel, Zendesk, IBM Research | 2026 |
| Member of free course/summer school's organizing committee | LxMLS 2025 - The 15th Lisbon Machine Learning School | IST; IT; INESC-ID; LUMLIS, Unbabel, Zendesk, IBM Research | 2025 |

| | | | |
|--|--|---|------|
| Member of free course/summer school's organizing committee | LxMLS 2024 - The 14th Lisbon Machine Learning School | IST; IT; INESC-ID; LUMLIS, Unbabel, Zendesk, IBM Research | 2024 |
| Member of free course/summer school's organizing committee | LxMLS 2023 - 13th Lisbon Machine Learning School | IST; IT; INESC-ID; LUMLIS, Unbabel, Zendesk, IBM Research | 2023 |
| Member of free course/summer school's organizing committee | LxMLS 2022 - 12th Lisbon Machine Learning School | IST; IT; INESC-ID; LUMLIS, Unbabel, Zendesk, IBM Research | 2022 |
| Member of free course/summer school's organizing committee | LxMLS 2021 - 11th Lisbon Machine Learning School | IT/IST/INESC-ID/Unbabel | 2021 |
| Member of free course/summer school's organizing committee | LxMLS 2020 - 10th Lisbon Machine Learning School | IT/IST/INESC-ID/Unbabel | 2020 |
| Member of free course/summer school's organizing committee | LxMLS 2019 - 9th Lisbon Machine Learning School | IT/IST/INESC-ID/Unbabel | 2019 |
| Member of free course/summer school's organizing committee | LxMLS 2018 - 8th Lisbon Machine Learning School | IT/L2F-INESC-ID | 2018 |
| Member of free course/summer school's organizing committee | LxMLS 2017 - 7th Lisbon Machine Learning School | IST/IT/Unbabel/INESC-ID Lisboa | 2017 |
| Member of free course/summer school's organizing committee | LxMLS 2016 - 6th Lisbon Machine Learning School | IST/IT/Unbabel/INESC-ID Lisboa | 2016 |
| Member of free course/summer school's organizing committee | LxMLS 2015 - 5th Lisbon Machine Learning School | IST/IT/Unbabel/INESC-ID Lisboa | 2015 |
| Member of scientific event's organizing committee | Machine Learning for NLP @ UALG | Universidade do Algarve | 2012 |