

PROPOR 2024

**Proceedings of the 16th International Conference on Computational
Processing of the Portuguese Language, PROPOR 2024 - vol. 1**

12–15 March, 2024
Universidade de Santiago de
Compostela, Galicia, Spain



©2024 The Association for Computational Linguistics

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 979-8-89176-062-2

Introduction

These Proceedings present the 16th edition of the International Conference on the Computational Processing of Portuguese (PROPOR 2024), held for the first time in Galicia, at the Universidade de Santiago de Compostela, Spain, during March 12–15, 2024.

PROPOR continues to be a key forum bringing together researchers dedicated to the computational processing of Portuguese, promoting the exchange of experiences in the development of methodologies, language resources, tools, applications, and innovative projects. In this edition, the conference was held in Galicia, the birthplace of the Portuguese language, and was extended to Galician (considered as a local variety of the former). PROPOR 2024 is, thus, the main scientific event in the area of natural language processing that is focused on theoretical and technological issues of written and spoken Portuguese and Galician.

This year’s edition includes an Industry track, alongside the General track. We received a total of 111 submissions — 81 long papers and 30 short papers —, from 152 different authors working on Portuguese and Galician around the world, in countries such as Portugal, Brazil, Spain, Norway, France, and Germany.

The Proceedings of PROPOR 2024 are divided in two volumes. The first volume presents the 51 long papers accepted (reflecting an acceptance rate of 63%), and the 20 short papers accepted (an acceptance rate of 66,6%), plus the papers presenting the two Best Dissertations, selected by the jury from a set of 12 candidates. All submissions were peer-reviewed by 104 reviewers from a total of 60 academic institutions and companies, coming from 14 countries. Full papers are organized thematically and show research and developments in resources and evaluation, natural language processing tasks, natural language processing applications, speech processing and applications, and lexical semantics. The second volume of the Proceedings is dedicated to the demos and to the works selected and presented in the four workshops hosted by PROPOR 2024, namely the *PROPOR’24 Competition on Automatic Essay Scoring of Portuguese Narrative Essays*, the *5th OpenCor: Latin American and Iberian Languages Open Corpora Forum*, the *Third Workshop on Digital Humanities and Natural Language Processing*, and the *First Workshop on NLP for Indigenous Languages of Lusophone Countries*.

The invited speakers of PROPOR 2024 are experts from different fields and with diverse research and professional experiences, showing the range, relevance and applicability of the computational processing of Portuguese and Galician:

- Elias Feijó, Lecturer of Portuguese language, literature, and culture at the Universidade de Santiago de Compostela and director of the Galabra Group, will reflect on the question “É este galego latim em pó?”.
- Gemma Boleda, ICREA Research Professor of the Department of Translation and Language Sciences at the Universitat Pompeu Fabra, will talk about ‘Pressures on the lexicon and their effects’.
- Marta Ruiz Costa-jussà, currently research scientist at FAIR, Meta, and recipient of an ERC Starting Grant and two Google Faculty Research Awards, will present “Beyond Semantic Evaluation in Seamless Speech Translation Models”.

PROPOR 2024 is a three and a half days conference that encompasses one full-day of workshops and shared tasks and two and a half days of communications, demos and community meetings.

Our sincere thanks go to every person and institution involved in the complex organization of this event, especially to the members of the Program Committee of the main event, the dissertations contest and the associated workshops chairs, the invited speakers, and the general organization staff. We are also grateful to the agencies and

organizations that supported and promoted the event.

Thank you all for participating and we wish you an enjoyable and inspiring conference!

Pablo Gamallo

Daniela Claro

António Teixeira

Livy Real

Marcos Garcia

Hugo Gonçalo Oliveira

Raquel Amaro

March 2024

Organization:

General Chairs: Pablo Gamallo and Daniela Claro

Program Chairs: António Teixeira, Livy Real, and Marcos Garcia

Editorial Chairs: Hugo Gonçalo Oliveira and Raquel Amaro

Demo Chairs: Iria de Dios and Marlo Souza

Workshop/Tutorial Chairs: Alberto Abad, Alberto Simões and Helena Caseli

Best Dissertation Chairs: António Branco, Paulo Quaresma and Renata Vieira

Industry Track Chairs: José Ramom Pichel and Luís Trigo

Local Organizers: Daniel Bardanca and José Ramom Pichel

Consultants: Carolina Scarton and Fernando Batista

Program Committee:

Adina Valdu, Universidade de Santiago de Compostela (Spain)

Adriana Pagano, Federal University of Minas Gerais (Brazil)

Alberto Abad, INESC-ID and Instituto Superior Técnico, Universidade de Lisboa (Portugal)

Alberto Simões, 2Ai Lab - IPCA (Portugal)

Alexandre Rademaker, IBM Research and FGV/EMAp (Brazil)

Aline Paes, Fluminense Federal University (Brazil)

Amália Mendes, University of Lisbon (Portugal)

Ana Isabel Mata, University of Lisbon (Portugal)

Ana Luísa V. Leal, University of Macau (Macau)

Antonio Bonafonte, Amazon (Spain)

António Branco, University of Lisbon (Portugal)

António Teixeira, University of Aveiro (Portugal)

Ariani Di Felippo, Federal University of São Carlos (Brazil)

Arnaldo Candido Junior, Federal University of Technology - Paraná (Brazil)

Berthold Crysmann, CNRS and University of Paris (France)

Brett Drury, Liverpool Hope University (UK)

Bruno Martins, INESC-ID and Instituto Superior Técnico, Universidade de Lisboa (Portugal)

Carlos A. Prolo, Federal University of Rio Grande do Norte (Brazil)

Carlos Ramisch, Aix-Marseille University (France)

Carmen Magariños, Universidade de Santiago de Compostela (Spain)

Catarina Oliveira, University of Aveiro (Portugal)

Christopher Shulby, Kin AI (Brazil)

Cristiane Namiuti, Universidade Estadual do Sudoeste da Bahia (Brazil)

Cristina Carbajal, Universidade de Santiago de Compostela (Spain)

Daniela Claro, Federal University of Bahia (Brazil)

Daniel Beck, University of Melbourne (Australia)

David Martins de Matos, INESC-ID and Instituto Superior Técnico, Universidade de Lisboa (Portugal)

Diamantino Freitas, University of Oporto (Portugal)

Diana Santos, Linguatca and University of Oslo (Norway)

Eloize Rossi M. Seno, Federal Institute of São Paulo (Brazil)

Eric Laporte, Université Gustave Eiffel (France)
Evelin Amorim, INESC TEC (Portugal)
Fernando Batista, INESC-ID and ISCTE-IUL (Portugal)
Fernando Perdigão, Instituto de Telecomunicações and University of Coimbra (Portugal)
Gaël Dias, University of Caen Normandy (France)
Helena Bermúdez, JinnTec GmbH (Germany)
Helena Caseli, Federal University of São Carlos (Brazil)
Hugo Gonçalo Oliveira, CISUC and University of Coimbra (Portugal)
Irene Rodrigues, University of Évora (Portugal)
Iria de Dios Flores, Universidade de Santiago de Compostela (Spain)
Isabel Falé, Universidade Aberta and University of Lisbon (Portugal)
Isabel Trancoso, INESC-ID and Instituto Superior Técnico, Universidade de Lisboa (Portugal)
Ivandré Paraboni, University of São Paulo (Brazil)
Ivanovitch Silva, Universidade Federal do Rio Grande do Norte (Brazil)
Javier González Corbelle, Universidade de Santiago de Compostela (Spain)
João Balsa, University of Lisbon (Portugal)
João Silva, University of Lisbon (Portugal)
Jorge Baptista, University of Algarve and INESC-ID (Portugal)
José Ramom Pichel, Universidade de Santiago de Compostela (Spain)
José Saias, University of Évora (Portugal)
Leandro Oliveira, Embrapa Informatica Agropecuaria (Spain)
Leonardo Zilio, Friedrich-Alexander-Universität Erlangen-Nürnberg (Germany)
Livy Real, QuintoAndar Inc. (Brazil)
Luciana Benotti, National University of Córdoba (Argentina)
Magali Duran, University of São Paulo (Brazil)
Marcelo Finger, University of São Paulo (Brazil)
Marcos Fernández Pichel, Universidade de Santiago de Compostela (Spain)
Marcos Garcia, Universidade de Santiago de Compostela (Spain)
Marcos Spalenza, UFES (Brazil)
Marcos Treviso, Instituto de Telecomunicações (Brazil)
Maria das Graças Volpes Nunes, University of São Paulo (Brazil)
Maria José B. Finatto, Federal University of Rio Grande do Sul (Brazil)
Mario Ezra Aragon, Universidade de Santiago de Compostela (Spain)
Mário Rodrigues, ESTGA/IEETA - University of Aveiro (Portugal)
Marlo Souza, Federal University of Bahia (Brazil)
Martín Pereira-Fariña, University of Santiago de Compostela (Spain)
Mikel L. Forcada, University of Alacant (Spain)
Nádia Silva, University of São Paulo (Brazil)
Nelson Neto, Federal University of Pará (Brazil)
Norton Roman, University of São Paulo (Brazil)
Oto Vale, Federal University of São Carlos (Brazil)
Pablo Faria, University of Campinas (Brazil)
Palmira Marrafa, University of Lisbon (retired) (Portugal)
Paula Cardoso, Federal University of Lavras (Brazil)
Paulo Quaresma, University of Évora (Portugal)
Plinio Barbosa, University of Campinas (Brazil)
Prakash Poudyal, Kathmandu University (Nepal)

Raquel Amaro, NOVA University Lisbon (Portugal)
Renata Vieira, CIDEHUS and University of Évora (Portugal)
Ricardo Ribeiro, INESC-ID and ISCTE (Portugal)
Ricardo Rodrigues, CISUC and Polytechnic Institute of Coimbra (Portugal)
Sandra Aluísio, University of São Paulo (Brazil)
Sara Mendes, University of Lisbon (Portugal)
Saullo Haniell, Pontifícia Universidade Católica de Campina (Brazil)
Sheila Castilho, Dublin City University (Ireland)
Sofía García, UPV/EHU (Basque Country/Spain)
Susana Duarte Martins, NOVA University Lisbon (Portugal)
Susana Sotelo Docío, University of Santiago de Compostela (Spain)
Tamás Gábor Csapó, Budapest University of Technology and Economics (Hungary)
Teresa Gonçalves, University of Évora (Portugal)
Thiago Ferreira, aiXplain inc. (USA)
Thiago Pardo, University of São Paulo (Brazil)
Valeria de Paiva, Topos Institute (USA)
Valeria Feltrim, Universidade Estadual de Maringá (Brazil)
Violeta Quental, Pontifical Catholic University of Rio de Janeiro (Brazil)
Vládía Pinheiro, Universidade de Fortaleza (Brazil)

Additional Reviewers:

Erik Bran Marino, Universidade de Évora (Portugal)
Márcio Lima Inácio, University of Coimbra (Portugal)
Isabel Carvalho, University of Coimbra (Portugal)
Rodrigo Wilkens, Université Catholique de Louvain (Belgium)
André Santos, INESC TEC (Portugal)
Felipe Ribas Serras, Universidade de São Paulo (Brazil)
Guido Ivorra, Universidad Nacional de Córdoba (Argentina)

Table of Contents

Long Papers

A Multilingual Dataset for Investigating Stereotypes and Negative Attitudes Towards Migrant Groups in Large Language Models	1
<i>Danielly Sorato, Carme Colominas Ventura and Diana Zavala-Rojas</i>	
Is ChatGPT an effective solver of sentiment analysis tasks in Portuguese? A Preliminary Study	13
<i>Gladson de Araujo, Tiago de Melo and Carlos Maurício S. Figueiredo</i>	
A Galician Corpus for Misogyny Detection Online	22
<i>Lucía M. Álvarez-Crespo and Laura M. Castro</i>	
Simple and Fast Automatic Prosodic Segmentation of Brazilian Portuguese Spontaneous Speech	32
<i>Giovana Meloni Craveiro, Vinicius Gonçalves Santos, Gabriel Jose Pellisser Dalalana, Flaviane R. Fernandes Svartman and Sandra Maria Aluísio</i>	
LLMs and Translation: different approaches to localization between Brazilian Portuguese and European Portuguese	45
<i>Eduardo G. Cortes, Ana Luíza Vianna, Mikaela Martins, Sandro Rigo and Rafael Kunst</i>	
SPARQL can also talk in Portuguese: answering natural language questions with knowledge graphs	56
<i>Elbe Miranda, Aline Paes and Daniel de Oliveira</i>	
Exploring Pre-Trained Transformers for Translating Portuguese Text to Brazilian Sign Language	67
<i>Jose Mario De Martino and Dener Stassun Christinele</i>	
NLP for historical Portuguese: Analysing 18th-century medical texts	76
<i>Leonardo Zilio, Rafaela Radünz Lazzari and Maria Jose Bocorny Finatto</i>	
Text Summarization and Temporal Learning Models Applied to Portuguese Fake News Detection in a Novel Brazilian Corpus Dataset	86
<i>Gabriel Lino Garcia, Pedro Henrique Paiola, Danilo Samuel Jodas, Luis Afonso Sugi and João Paulo Papa</i>	
Automatic Text Readability Assessment in European Portuguese	97
<i>Eugénio Ribeiro, Nuno Mamede and Jorge Baptista</i>	
Toxic Speech Detection in Portuguese: A Comparative Study of Large Language Models	108
<i>Amanda da Silva Oliveira, Thiago de Carvalho Cecote, João Paulo Reis Alvarenga, Vander Luis de Souza Freitas and Eduardo José da Silva Luz</i>	
Named entity recognition specialised for Portuguese 18th-century History research	117
<i>Joaquim Santos, Helena Freire Cameron, Fernanda Olival, Fátima Farrica and Renata Vieira</i>	
Exploring Open Information Extraction for Portuguese Using Large Language Models	127
<i>Bruno Cabral, Daniela Claro and Marlo Souza</i>	
Bringing Pragmatics to Porttinari - Adding Speech Acts to News Texts	137
<i>Nataly L. Patti da Silva, Norton Trevisan Roman and Ariani Di Felippo</i>	
Authorship Attribution with Rejection Capability in Challenging Contexts of Limited Datasets	146
<i>Pedro Oliveira and Joaquim Silva</i>	
Using Large Language Models for Identifying Satirical News in Brazilian Portuguese	156
<i>Gabriela Wick-Pedro, Cássio Faria da Silva, Marcio Lima Inácio, Oto Araújo Vale and Helena de Medeiros Caseli</i>	
Semantic Permanence in Audiovisual Translation: a FrameNet approach to subtitling	168

<i>Mairon Samagaio, Tiago Torrent, Ely Matos and Arthur Almeida</i>	
Hurdles in Parsing Multi-word Adverbs: Examples from Portuguese	177
<i>Izabela Muller, Nuno Mamede and Jorge Baptista</i>	
Portal NURC-SP: Design, Development, and Speech Processing Corpora Resources to Support the Public Dissemination of Portuguese Spoken Language	187
<i>Ana Carolina Rodrigues, Alessandra A. Macedo, Arnaldo Candido Jr, Flaviane R. F. Svartman, Giovana M. Craveiro, Marli Quadros Leite, Sandra M. Aluísio, Vinícius G. Santos and Vinícius M. Garcia</i>	
TransAlign: An Automated Corpus Generation through Cross-Linguistic Data Alignment for Open Information Extraction	196
<i>Alan Rios, Bruno Cabral, Daniela Claro, Rerisson Cavalcante and Marlo Souza</i>	
BATS-PT: Assessing Portuguese Masked Language Models in Lexico-Semantic Analogy Solving and Relation Completion	207
<i>Hugo Gonçalo Oliveira, Ricardo Rodrigues, Bruno Ferreira, Purificação Silvano and Sara Carvalho</i>	
Towards the automatic creation of NER systems for new domains	218
<i>Emanuel Matos, Mário Rodrigues and António Teixeira</i>	
A New Benchmark for Automatic Essay Scoring in Portuguese	228
<i>Igor Cataneo Silveira, André Barbosa and Denis Deratani Mauá</i>	
Predicting the Age of Emergence of Consonants	238
<i>Luís Jesus and Jihen Trabelsi</i>	
Applying event classification to reveal the Estado da Índia	247
<i>Gonçalo C. Albuquerque, Marlo Souza, Renata Vieira and Ana Sofia Ribeiro</i>	
Exploring Computational Discernibility of Discourse Domains in Brazilian Portuguese within the Carolina Corpus	255
<i>Felipe Ribas Serras, Mariana Sturzeneker, Miguel de Mello Carpi, Mayara Feliciano Palma, Maria Clara Ramos Morales Crespo, Aline Silva Costa, Vanessa Martins Do Monte, Cristiane Namiuti, Maria Clara Paixão de Souza and Marcelo Finger</i>	
Identification of Types of Event-Time Temporal Relations in Portuguese Using a Rule-Based Approach	266
<i>Dárcio S. Rocha, Marlo Souza and Daniela B. Claro</i>	
A Corpus of Stock Market Tweets Annotated with Named Entities	276
<i>Michel Monteiro Zerbinati, Norton Trevisan Roman and Ariani Di Felippo</i>	
Frequency, overlap and origins of palatal sonorants in three Iberian languages	285
<i>Carlos Silva and Luís Trigo</i>	
A Named Entity Recognition Approach for Portuguese Legislative Texts Using Self-Learning	290
<i>Rafael Oleques Nunes, Dennis Giovani Balreira, André Suslik Spritzer and Carla Maria Dal Sasso Freitas</i>	
Exploring Portuguese Hate Speech Detection in Low-Resource Settings: Lightly Tuning Encoder Models or In-Context Learning of Large Models?	301
<i>Gabriel Assis, Annie Amorim, Jonnatahn Carvalho, Daniel de Oliveira, Daniela Vianna and Aline Paes</i>	
Aspect-based sentiment analysis in comments on political debates in Portuguese: evaluating the potential of ChatGPT	312
<i>Eloize Seno, Lucas Silva, Fábio Anno, Fabiano Rocha and Helena Caseli</i>	
CLSJR.BR - A Model for Abstractive Summarization of Legal Documents in Portuguese Language based on Contrastive Learning	321

Alex Aguiar Lins, Cecilia Silvestre Carvalho, Francisco Das Chagas Jucá Bomfim, Daniel de Carvalho Bentes and Vlândia Pinheiro

Investigating the Generalizability of Portuguese Readability Assessment Models Trained Using Linguistic Complexity Features	332
<i>Soroosh Akef, Amália Mendes, Detmar Meurers and Patrick Rebuschat</i>	
UlyssesNERQ: Expanding Queries from Brazilian Portuguese Legislative Documents through Named Entity Recognition	342
<i>Hidelberg Albuquerque, Ellen Souza, Tainan Silva, Rafael P. Gouveia, Flavio Junior, Douglas Vitório, Nádia F. F. da Silva, André C.P.L.F. de Carvalho, Adriano L.I. Oliveira and Francisco Edmundo de Andrade</i>	
Across the Atlantic: Distinguishing Between European and Brazilian Portuguese Dialects	353
<i>David Preda, Tomás Osório and Henrique Lopes Cardoso</i>	
Accent Classification is Challenging but Pre-training Helps: a case study with novel Brazilian Portuguese datasets	364
<i>Ariadne Matos, Gustavo Araújo, Arnaldo Candido Junior and Moacir Ponti</i>	
RoBERTaLexPT: A Legal RoBERTa Model pretrained with deduplication for Portuguese	374
<i>Eduardo A. S. Garcia, Nadia F. F. Silva, Felipe Siqueira, Hidelberg O. Albuquerque, Juliana R. S. Gomes, Ellen Souza and Eliomar A. Lima</i>	
Evaluating Pre-training Strategies for Literary Named Entity Recognition in Portuguese	384
<i>Mariana O. Silva and Mirella M. Moro</i>	
Brazilian Portuguese Product Reviews Moderation with AutoML	394
<i>Lucas Nildaimon dos Santos Silva, Livy Real, Fernando Rezende Zagatti, Ana Claudia Bianchini Zandavalle, Tatiana da Silva Gama and Carolina Francisco Gadelha Rodrigues</i>	
Towards Portparser - a highly accurate parsing system for Brazilian Portuguese following the Universal Dependencies framework	401
<i>Lucelene Lopes and Thiago Pardo</i>	
Study of the State of the Art Galician Machine Translation: English-Galician and Spanish-Galician models	411
<i>Sofía García González and German Rigau Claramunt</i>	
Bartoli's areal norms revisited: an agent-based modeling approach	422
<i>Dalmo Buzato and Evandro Cunha</i>	
RePro: a benchmark for Opinion Mining for Brazilian Portuguese	432
<i>Lucas Nildaimon dos Santos Silva, Livy Real, Ana Claudia Bianchini Zandavalle, Carolina Francisco Gadelha Rodrigues, Tatiana da Silva Gama, Fernando Guedes Souza and Phillipe Derwich Silva Zaidan</i>	
Glória: A Generative and Open Large Language Model for Portuguese	441
<i>Ricardo Lopes, Joao Magalhaes and David Semedo</i>	
Evaluating large language models for the tasks of PoS tagging within the Universal Dependency framework	454
<i>Mateus Machado and Evandro Ruiz</i>	
Question Answering for Dialogue State Tracking in Portuguese	461
<i>Francisco Pais, Patricia Ferreira, Catarina Silva, Ana Alves and Hugo Gonçalo Oliveira</i>	
Toxic Content Detection in online social networks: a new dataset from Brazilian Reddit Communities	472
<i>Luiz Henrique Quevedo Lima, Adriana Silvina Pagano and Ana Paula Couto da Silva</i>	
A Natural Language Text to Role-Playing Game Animation Generator	483
<i>Caio F. Oliveira, Artur Franco, Wellington Franco and José G. R. Maia</i>	

From Random to Informed Data Selection: A Diversity-Based Approach to Optimize Human Annotation and Few-Shot Learning	492
<i>Alexandre Alcoforado, Lucas Hideki Takeuchi Okamura, Israel Campos Fama, Bárbara Fernandes Dias Bueno, Arnold Moya Lavado, Thomas Palmeira Ferraz, Bruno Veloso and Anna Helena Reali Costa</i>	
Enhancing Stance Detection in Low-Resource Brazilian Portuguese Using Corpus Expansion generated by GPT-3.5	503
<i>Dyonnatan Maia and Nádia Félix Felipe da Silva</i>	
Short Papers	
A Bag-of-Users approach to mental health prediction from social media data	509
<i>Rafael Oliveira and Ivandré Paraboni</i>	
Semi-automatic corpus expansion: the case of stance prediction	515
<i>Camila Pereira and Ivandré Paraboni</i>	
Sequence-to-sequence and transformer approaches to Portuguese text style transfer	521
<i>Pablo Costa and Ivandré Paraboni</i>	
Comparative Analysis of Intentional Gramatical Error Correction Techniques on Twitter/X	527
<i>Thainá Marini and Taffarel Brant-Ribeiro</i>	
Towards a Syntactic Lexicon of Brazilian Portuguese Adjectives	532
<i>Ryan Martinez, Jorge Baptista and Oto Vale</i>	
Literary similarity of novels in Portuguese	539
<i>Diana Santos</i>	
An evaluation of Portuguese language models' adaptation to African Portuguese varieties	544
<i>Diego Fernando Válio Antunes Alves</i>	
Text Readability Assessment in European Portuguese: A Comparison of Classification and Regression Approaches	551
<i>Eugénio Ribeiro, Nuno Mamede and Jorge Baptista</i>	
Is it safe to machine translate suicide-related language from English to Galician?	558
<i>John E. Ortega and Annika Marie Schoene</i>	
First assessment of Graph Machine Learning approaches to Portuguese Named Entity Recognition	563
<i>Gabriel Silva, Mário Rodrigues, António Teixeira and Marlene Amorim</i>	
Exploring Multimodal Models for Humor Recognition in Portuguese	568
<i>Marcio Inácio and Hugo Gonçalo Oliveira</i>	
RecognaSumm: A Novel Brazilian Summarization Dataset	575
<i>Pedro Henrique Paiola, Gabriel Lino Garcia, Danilo Samuel Jodas, João Vitor Mariano Correia, Luis Afonso Sugi and João Paulo Papa</i>	
A Speech-Driven Talking Head based on a Two-Stage Generative Framework	580
<i>Brayan Bernardo and Paula Costa</i>	
Increasing manually annotated resources for Galician: the Parallel Universal Dependencies Treebank	587
<i>Xulia Sánchez-Rodríguez, Albina Sarymsakova, Laura Castro and Marcos Garcia</i>	
CorpusNÓS: A massive Galician corpus for training large language models	593
<i>Iria de-Dios-Flores, Silvia Paniagua Suárez, Cristina Carbajal Pérez, Daniel Bardanca Outeiriño, Marcos Garcia and Pablo Gamallo</i>	

Exploring the effects of vocabulary size in neural machine translation: Galician as a target language	600
<i>Daniel Bardanca Outeirinho, Pablo Gamallo Otero, Iria de-Dios-Flores and José Ramom Pichel Campos</i>	
A Reproducibility Analysis of Portuguese Computational Processing Conferences: A Case of Study	605
<i>Daniel Leal, Anthony Luz and Rafael Anchiêta</i>	
Automated admissibility of complaints about fraud and corruption	610
<i>Thiago De Paula, André Do Amaral, Andre Victor, Luis Alberto Sales, Rodrigo Moreira, Thiago Meirelles and Rafael Basso</i>	
Natural Language Processing Application in Legislative Activity: a Case Study of Similar Amendments in the Brazilian Senate	614
<i>Diany Pressato, Pedro Lucas Castro de Andrade, Flávio Rocha Junior, Felipe Alves Siqueira, Ellen Polliana Ramos Souza, Nádia Félix Felipe da Silva, Márcio de Souza Dias and André Carlos Ponce de Leon Ferreira de Carvalho</i>	
Spatial Information Challenges in English to Portuguese Machine Translation	620
<i>Rafael Fernandes, Rodrigo Souza, Marcos Lopes, Paulo Santos and Thomas Finbow</i>	
Compilation and tagging of a corpus with Celpe-Bras texts	627
<i>Juliana Schoffen, Elisa Stumpf, Deise Amaral, Luiza Divino, Isadora Hanauer, Isabel Lisboa, Amanda Raupp and Brenda Xavier</i>	
Best Dissertations	
TTS applied to the generation of datasets for automatic speech recognition	633
<i>Edresson Casanova, Sandra Aluísio and Moacir Antonelli Ponti</i>	
Text clustering applied to unbalanced data in legal contexts (Final Version)	639
<i>Lucas José Gonçalves Freitas</i>	