

Volume 8, Número 2 – Dezembro 2016

# LinguaMÁTICA

ISSN: 1647-0818

## **Editores ASSIN**

---

*Erick Fonseca*

*Leandro Santos*

*Marcelo Criscuolo*

*Sandra Aluísio*

## **Editores**

---

*Alberto Simões*

*José João Almeida*

*Xavier Gómez Guinovart*



# Conteúdo

## Avaliação de Similaridade Semântica e de Inferência Textual

### Visão Geral da ASSIN

*Erick Fonseca, Leandro dos Santos, Marcelo Criscuolo & Sandra Alúcio . . . . .* 3

### Blue Man Group no ASSIN: Usando Representações Distribuídas para Similaridade Semântica e Inferência Textual

*Luciano Barbosa, Paulo Cavalin, Victor Guimarães & Matthias Kormaksson . . . . .* 15

### FlexSTS: Um Framework para Similaridade Semântica Textual

*Jânio Freire, Vlândia Pinheiro & David Feitosa . . . . .* 23

### INESC-ID@ASSIN: Medição de Similaridade Semântica e Reconhecimento de Inferência Textual

*Pedro Fialho, Ricardo Marques, Bruno Martins, Luísa Coheur & Paulo Quaresma . . . . .* 33

### ASAPP: Alinhamento Semântico Automático de Palavras aplicado ao Português

*Ana Oliveira Alves, Ricardo Rodrigues & Hugo Gonçalo Oliveira . . . . .* 43

### Solo Queue at ASSIN: Combinando Abordagens Tradicionais e Emergentes

*Nathan Siegle Hartmann . . . . .* 59



# Editorial

*Neste oitavo ano de vida a Linguamática teve, como é seu hábito, duas edições. A primeira, em Julho, apenas com três artigos, e esta edição, de Dezembro, com artigos alargados correspondentes a uma workshop realizada com conjunto sim a conferência PROPOR'2016, a ASSIN: Avaliação de Similaridade Semântica e de Inferência Textual. Se por um lado o número de artigos publicado é pequeno, por outro, o número de artigos recebidos para avaliação não o foi, estando dentro da média habitual da Linguamática.*

*Mas o ano de 2016, embora mau, como sabemos, para muitas personalidades do mundo da música, não o foi para a Linguamática, que continua a ser indexada pela Scopus (embora o site da Scopus ainda não inclua todas as edições recentes), e passou a ser incluída na Web of Science da Thomson Reuters, no índice ESCI (Emerging Sources Citation Index), um índice de revistas selecionadas para avaliação e possível integração nos índices de topo. São, sem dúvidas, duas notícias que nos fazem muito orgulhosos dos nossos autores.*

*Xavier Gómez Guinovart*

*José João Almeida*

*Alberto Simões*



# Prólogo

## Avaliação de Similaridade Semântica e de Inferência Textual

*A Avaliação de Similaridade Semântica e de Inferência Textual (ASSIN) foi proposta como um Workshop em paralelo com o PROPOR 2016 para apresentação dos resultados da avaliação conjunta de duas subtarefas relacionadas, tratando da língua portuguesa, especificamente do Português do Brasil (PB) e Europeu (PE). Ambas as subtarefas dizem respeito ao entendimento de um par de sentenças: a similaridade semântica (STS, Semantic Textual Similarity) é uma medida numérica de 1 a 5 do quão similar é o conteúdo das duas sentenças; e a inferência textual (RTE, Recognizing Textual Entailment) consiste em classificar o par como tendo uma relação de implicação, paráfrase, ou nenhuma das duas.*

*A avaliação conjunta deixou como legado o córpus ASSIN de 10.000 pares de sentenças (5.000 em PB e 5.000 em PE) usado pelos participantes e que está publicamente disponível em <http://nilc.icmc.usp.br/assin/>. Somos gratos a todos os anotadores do córpus, pois sem eles a avaliação não teria sido realizada.*

*Foram seis os participantes da avaliação: três do Brasil (Solo Queue, Blue Man Group, LEC-UNIFOR) e três de Portugal (INESC-ID, ASAPP, Reciclagem) sendo que todos participaram da tarefa STS, e somente quatro deles da tarefa RTE.*

*Nesta edição especial da Linguamática em homenagem ao Workshop ASSIN, trazemos o artigo com a apresentação da Avaliação Conjunta e mais cinco versões revisadas e estendidas dos seguintes artigos apresentados no Workshop, sendo que as equipes ASAPP e Reciclagem escreveram um único artigo reportando ambos os resultados.*

*Desejamos a todos uma leitura proveitosa destes trabalhos!*

*Erick Fonseca  
Leandro Santos  
Marcelo Criscuolo  
Sandra Aluísio*



# Comissão Científica

**Alberto Álvarez Lugrís,**  
Universidade de Vigo

**Alberto Simões,**  
Universidade do Minho

**Aline Villavicencio,**  
Universidade Federal do Rio Grande do Sul

**Álvaro Iriarte Sanroman,**  
Universidade do Minho

**Ana Frankenberg-Garcia,**  
University of Surrey

**Anselmo Peñas,**  
Univers. Nac. de Educación a Distancia

**Antón Santamarina,**  
Universidade de Santiago de Compostela

**Antoni Oliver González,**  
Universitat Oberta de Catalunya,

**Antonio Moreno Sandoval,**  
Universidad Autónoma de Madrid

**António Teixeira,**  
Universidade de Aveiro

**Arantza Díaz de Ilarraza,**  
Euskal Herriko Unibertsitatea

**Arkaitz Zubiaga,**  
Dublin Institute of Technology

**Belinda Maia,**  
Universidade do Porto

**Carmen García Mateo,**  
Universidade de Vigo

**Diana Santos,**  
Linguatca/Universidade de Oslo

**Ferran Pla,**  
Universitat Politècnica de València

**Gael Harry Dias,**  
Université de Caen Basse-Normandie

**Gerardo Sierra,**  
Univers. Nacional Autónoma de México

**German Rigau,**  
Euskal Herriko Unibertsitatea

**Helena de Medeiros Caseli,**  
Universidade Federal de São Carlos

**Horacio Saggion,**  
University of Sheffield

**Hugo Gonçalo Oliveira,**  
Universidade de Coimbra

**Iñaki Alegria,**  
Euskal Herriko Unibertsitatea

**Irene Castellón Masalles,**  
Universitat de Barcelona

**Joaquim Llisterri,**  
Universitat Autònoma de Barcelona

**José João Almeida,**  
Universidade do Minho

**José Paulo Leal,**  
Universidade do Porto

**Joseba Abaitua,**  
Universidad de Deusto

**Juan-Manuel Torres-Moreno,**  
Lab. Informatique d'Avignon - UAPV

**Kepa Sarasola,**  
Euskal Herriko Unibertsitatea

**Laura Plaza,**  
Complutense University of Madrid

**Lluís Padró,**  
Universitat Politècnica de Catalunya

**Marcos Garcia,**  
Universidade de Santiago de Compostela

**María Inés Torres,**  
Euskal Herriko Unibertsitatea

**Maria das Graças Volpe Nunes,**  
Universidade de São Paulo

**Mercè Lorente Casafont,**  
Universitat Pompeu Fabra

**Mikel Forcada,**  
Universitat d'Alacant

**Pablo Gamallo Otero,**  
Universidade de Santiago de Compostela

**Patrícia Cunha França,**  
Universidade do Minho

**Rui Pedro Marques,**  
Universidade de Lisboa

**Salvador Climent Roca,**  
Universitat Oberta de Catalunya

**Susana Afonso Cavadas,**  
University of Sheffield

**Tony Berber Sardinha,**  
Pontifícia Univ. Católica de São Paulo

**Xavier Gómez Guinovart,**  
Universidade de Vigo