

Volume 6, Número 1 – Julho 2014

LinguaMÁTICA

ISSN: 1647-0818

Editoras STIL

Sandra Maria Aluísio

Valéria Delisandra Feltrim

Editores

Alberto Simões

José João Almeida

Xavier Gómez Guinovart

Conteúdo

Simpósio de Tecnologia da Informação e Linguagem Humana

Geração de expressões de referência em ambientes virtuais interativos <i>Diego dos Santos Silva e Ivandré Paraboni</i>	15
Usando grades de entidades na análise automática de coerência local em textos científicos <i>Alison Rafael Polpeta Freitas e Valéria Delisandra Feltrim</i>	29
NERP-CRP: uma ferramenta para o reconhecimento de entidades nomeadas por meio de Conditional Random Fields <i>Daniela Oliveira F. do Amaral e Renata Vieira</i>	41

Artigos de Investigação

Realização de previsões com conteúdos textuais em Português <i>Indira Mascarenhas Brito e Bruno Martins</i>	53
PoNTE: apontando para corpos de aprendizes de tradução avançados <i>Diana Santos</i>	69

Editorial

Este ano de 2014 é iniciado com uma edição especial. Assim como em 2010, publicamos um conjunto de artigos alargados, seleccionados dos artigos aceites no nono Simpósio Brasileiro de Tecnologia da Informação e Linguagem Humana (STIL).

Portanto, esta edição abre com três artigos seleccionados da edição de 2013 do STIL que abordam diferentes aspectos da linguagem natural: em primeiro lugar a geração de texto de forma a descrever o ambiente virtual em que um utilizador se encontra; posteriormente será discutida a análise de coerência no uso de entidades em textos científicos; finalmente será apresentado um sistema para o reconhecimento de entidades mencionadas, ou nomeadas.

Para completar o volume, incluímos neste número especial dois artigos de investigação que não fazem parte do STIL: primeiro um trabalho relacionado com a previsão, usando diferentes tipos de regressão e dados extraídos de texto escrito em linguagem natural; e em seguida, um trabalho sobre a anotação e disponibilização de corpos paralelos criados a partir de trabalhos de tradução de alunos, de modo a serem úteis para, entre outras coisas, o próprio ensino de línguas.

Xavier Gómez Guinovart

José João Almeida

Alberto Simões

Prólogo

Uma visão geral dos avanços no Simpósio de Tecnologia da Informação e Linguagem Humana

O Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana (STIL) é o principal evento nacional apoiado e organizado pela Comissão Especial de Processamento de Linguagem Natural (CE-PLN¹) da Sociedade Brasileira de Computação (SBC²).

O evento foi concebido em 2003 com o nome TIL (Workshop de Tecnologia da Informação e da Linguagem Humana), tendo o propósito de estimular o desenvolvimento de uma área genuinamente multidisciplinar, procurando atrair pesquisadores, membros da comunidade acadêmica e da indústria que atuam nas áreas de Ciência da Computação, Linguística e Ciência da Informação, entre outras, pois o processamento computacional das línguas humanas requer a coordenação de esforços de diversas comunidades, que contribuem com conhecimentos específicos e metodologias de pesquisa próprias no desenvolvimento de técnicas e sistemas. O principal objetivo do STIL é fornecer o fórum adequado para a integração dessas comunidades.

Em 2003, foi realizado na USP-São Carlos/SP; em 2004 e 2005 foi hospedado pelo Congresso da SBC em Salvador/BA e São Leopoldo/RS, respectivamente; em 2006, o evento foi hospedado pela International Joint Conference IBERAMIA/SBIA/SBRN, em Ribeirão Preto/SP, que consistiu no maior evento de Inteligência Artificial já realizado no Brasil. A 5a. edição do evento foi hospedada novamente pelo XXVII Congresso da SBC no Rio de Janeiro/RJ, no Instituto Militar de Engenharia-IME. A 6a. edição do evento foi realizado em 2008 juntamente com o Webmedia, em Vila Velha-ES, e foi a última com o nome de TIL. A 7a. edição, já com o nome de STIL, foi realizada na USP-São Carlos/SP em 2009. A 8a edição ocorreu em Cuiabá/MT em 2011, na UFMT. A 9a. edição foi realizada em Fortaleza/CE em 2013, juntamente com o 2o. Brazilian Conference on Intelligent Systems (BRACIS-13) e o X Encontro Nacional de Inteligência Artificial e Computacional (ENIAC). Atualmente, o STIL é o maior evento da comunidade no Brasil, podendo ser considerado o único no país totalmente dedicado ao tema.

A edição de 2013 recebeu 65 submissões do Brasil, Grã-Bretanha, Peru, Alemanha, Estados Unidos e Portugal. Cada artigo foi revisado por, pelo menos, 3 membros do Comitê de Programa que contava com 64 membros de 7 países e 34 instituições de ensino superior.

¹<http://www.nilc.icmc.usp.br/cepln/>

²<http://www.sbc.org.br/>

15 artigos foram selecionados para apresentação oral (taxa de aceitação de 23%) e 17 para apresentação na forma de pôster. Os melhores trabalhos foram convidados para apresentar versões estendidas a serem publicadas em duas revistas importantes da área da Computação e do Processamento Computacional das Línguas Humanas, respectivamente, o Journal of the Brazilian Computer Society (JBCS) e a Linguamática – Revista para o Processamento Automático das Línguas Ibéricas.

Nesta edição especial da Linguamática em homenagem ao STIL 2013, trazemos três versões estendidas dos seguintes artigos publicados na edição de 2013 do STIL:

- Geração de expressões de referência em ambientes virtuais interativos, por *Diego Silva e Ivandré Paraboni*;
- Usando grades de entidades na análise automática de coerência local em textos científicos, por *Alison Polpeta Freitas e Valéria Feltrim*;
- NERP-CRF: uma ferramenta para o reconhecimento de entidades nomeadas por meio de conditional random fields, por *Daniela do Amaral e Renata Vieira*;

Aproveitamos a oportunidade para agradecer aos autores, membros do Comitê de Programa, Palestrantes Convidados, SBC e os comitês locais e gerais do STIL 2013.

Desejamos a todos uma leitura proveitosa destes trabalhos!

*Sandra Maria Aluísio
Valéria Delisandra Feltrim*

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,
Universidade Beira Interior

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 Carlos Medeiros,
Porto Editora

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

Lluís Padró,
Universitat Politècnica de Catalunya

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