

# **X Brazilian Symposium in Information and Human Language Technology and Collocated Events (STIL 2015)**

Natal, Brazil  
4 – 7 November 2015

**Editors:**

**Claudia Freitas  
Alexandre Rademaker**

ISBN: 978-1-5108-1782-1

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2015) by the Association for Computational Linguistics  
All rights reserved.

Printed by Curran Associates, Inc. (2015)

For permission requests, please contact the Association for Computational Linguistics  
at the address below.

Association for Computational Linguistics  
209 N. Eighth Street  
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006  
Fax: 1-570-476-0860

[acl@aclweb.org](mailto:acl@aclweb.org)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2634  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# Contents

## I Conference Papers

### 1 Invited Talks

Numeric and Symbolic NLP: A Promising Engagement <i>Adam Pease</i> . . . . .	N/A
FreeLing: All you wanted to know and were afraid to ask <i>Lluís Padró</i> . . . . .	N/A

### 2 Short Papers 14

JCLex: A Java Tool for Compiling Finite-State Transducers from Full-Form Lexicons <i>Leonel F. de Alencar and Philipp B. Costa and Mardônio J. C. França and Alexander Ewart and Katuscia M. Andrade and Rossana M. C. Andrade</i> . . . . .	15
7x1-PT: um Corpus extraído do Twitter para Análise de Sentimentos em Língua Portuguesa <i>Silvia M. W. Moraes and Isabel H. Manssour and Milene S. Silveira</i> . . . . .	21
Comparative Analysis between Notations to Classify Named Entities using Conditional Random Fields <i>Daniela Oliveira F. do Amaral and Maiki Buffet and Renata Vieira</i> . . . . .	27
Do Extrator de Conhecimento Coletivo à Ágora Virtual: desenvolvendo uma ferramenta para democracia participativa <i>Tiago Novaes Angelo and César José Bonjuani Pagan and Romis Ribeiro Faissol Attux and Ricardo Ribeiro Gudwin</i> . . . . .	33
Um novo corpo e os seus desafios <i>Diana Santos</i> . . . . .	39
Análise Automática de Coerência Textual em Resumos Científicos: Avaliando Quebras de Linearidade <i>Leandro Lago da Silva and Valéria Delisandra Feltrim</i> . . . . .	45
Anotação de corpus com a OpenWordNet-PT: um exercício de desambiguação <i>Cláudia Freitas and Livy Real and Alexandre Rademaker</i> . . . . .	51
Integrating support verb constructions into a parser <i>Amanda Rassi and Jorge Baptista and Nuno Mamede and Oto Vale</i> . . . . .	57

## CONTENTS

---

Extração de Alvos em Comentários de Notícias em Português baseada na Teoria da Centralização <i>Frank Willian Cardoso de Oliveira and Valéria Delisandra Feltrim</i> . . . . .	63
Portal Min@s: Uma Ferramenta Geral de Apoio ao Processamento de Córpus de Propósito Geral <i>Arnaldo Candido Junior and Thiago Lima Vieira and Marcel Serikawa and Matheus Antônio Ribeiro Silva and Régis Zangirolami and Sandra Maria Aluísio</i> . . . . .	69
PrepNet.Br: a Semantic Network for Prepositions. <i>Débora D. Garcia and Bento Carlos Dias da Silva</i> . . . . .	75
<b>3 Full Papers</b>	<b>80</b>
Joint semantic discourse models for automatic multi-document summarization <i>Paula C. Figueira Cardoso and Thiago A. S. Pardo</i> . . . . .	81
Building and Applying Profiles Through Term Extraction <i>Lucelene Lopes and Renata Vieira</i> . . . . .	91
An Annotated Corpus for Sentiment Analysis in Political News <i>Gabriel Domingos de Arruda and Norton Trevisan Roman and Ana Maria Monteiro</i>	101
Campos Aleatórios Condicionais Aplicados à Detecção de Estrutura Retórica em Resumos de Textos Acadêmicos em Português <i>Alexandre C. Andreani and Valéria D. Feltrim</i> . . . . .	111
Anotando um Corpus de Notícias para a Análise de Sentimentos: um Relato de Experiência <i>Mariza Miola Dosciatti and Lohann Paterno Coutinho Ferreira and Emerson Cabrera Paraiso</i> . . . . .	121
Tesauros Distribucionais para o Português: avaliação de metodologias <i>Rodrigo Wilkens and Leonardo Zilio and Eduardo Ferreira and Gabriel Gonçalves and Aline Villavicencio</i> . . . . .	131
On Strategies of Human Multi-Document Summarization <i>Renata Tironi de Camargo and Ariani Di Felippo and Thiago A. S. Pardo</i> . . . . .	141
Enriching entity grids and graphs with discourse relations: the impact in local coherence evaluation <i>Márcio de S. Dias and Thiago A. S. Pardo</i> . . . . .	151
VerbLexPor: um recurso léxico com anotação de papéis semânticos para o português <i>Leonardo Zilio and Maria José Bocorny Finatto and Aline Villavicencio</i> . . . . .	161
Novo dicionário de formas flexionadas do Unitex-PB: avaliação da flexão verbal <i>Oto A. Vale and Jorge Baptista</i> . . . . .	171
Desambiguação de Homógrafos-Heterófonos por Aprendizado de Máquina em Português Brasileiro <i>Leonardo Hamada and Nelson Neto</i> . . . . .	181
RePort - Um Sistema de Extração de Informações Aberta para Língua Portuguesa <i>Victor Pereira and Vlória Pinheiro</i> . . . . .	191

## CONTENTS

---

Semi-Automatic Construction of a Textual Entailment Dataset: Selecting Candidates with Vector Space Models <i>Erick R. Fonseca and Sandra Maria Aluísio</i> . . . . .	201
n-Gramas de Caractere como Técnica de Normalização Morfológica para Língua Portuguesa: Um Estudo em Categorização de Textos <i>Guilherme T. Guimarães and Marcus V. Meirose and Silvia M. W. Moraes</i> . . . . .	211

## II IV Jornada de Descrição do Português

### 4 Apresentação Oral

A utilização de atos de diálogo em sistemas de diálogo para dispositivos móveis. <i>Tiago Martins da Cunha and Daniel de França Brasil Soares</i> . . . . .	N/A
Análise contrastiva da classificação sintático-semântica dos verbos locativos no Português do Brasil e no Português europeu <i>Roana Rodrigues and Jorge Baptista and Oto Vale</i> . . . . .	N/A
A criação de um corpus de sentenças através de gramáticas livres de contexto <i>Tiago Martins da Cunha and Paulo Bruno Lopes da Silva</i> . . . . .	N/A
Em direção à caracterização da complementaridade no corpus multidocumento CST-News <i>Jackson Souza and Ariani Di Felippo</i> . . . . .	N/A
Explorando hierarquias conceituais para a seleção de conteúdo na sumarização automática multidocumento <i>Andressa C. I. Zacarias and Ariani Di Felippo</i> . . . . .	N/A
Importância dos falsos homógrafos para a correção automática de erros ortográficos em Português <i>Magali Sanches Duran and Lucas Vinícius Avanço and Maria das Graças Volpe Nunes</i> . . . . .	N/A
A inconsistência do tratamento dispensado às preposições pela gramática tradicional <i>Débora Domiciano Garcia and Bento Carlos Dias da Silva</i> . . . . .	N/A

### 5 Apresentação por Pôster

As estratégias linguísticas e cognitivas que regem o Internetês - a escrita em rede - nos comentários do Facebook <i>Cristina Normandia and Maria Teresa Tedesco V. Abreu</i> . . . . .	N/A
Estudos recentes sobre a detecção de contradição no processamento automático de línguas naturais <i>Denis Luiz Marcello Owa</i> . . . . .	N/A
Transdutor de estados finitos para reconhecimento da nasalidade na pronúncia da variedade potiguar <i>Cid Ivan da Costa Carvalho</i> . . . . .	N/A