

CURRICULUM VITAE

INFORMAÇÕES PESSOAIS

Nome: Gustavo Henrique Paetzold

Nascido em: 13/09/1991

Telefone: +44 7476690262

E-mail: ghpaezold@outlook.com
g.h.paezold@sheffield.ac.uk

Website: <https://gustavopaetzold.wordpress.com>

Github: <https://github.com/ghpaezold>

INTERESSES PRINCIPAIS

Aprendizado de Máquina

- Aprendizado profundo (*deep learning*)
- Redes neurais convolucionais e recorrentes
- Modelos supervisionados de ranqueamento
- Modelos dirigidos por dados (*data-driven learning*)

Processamento de Linguagem Natural

- Adaptação de texto
- Simplificação de texto
- Estimativa de qualidade para tradução
- Algoritmos de transdução de árvores

Métodos de Compilação

- Análise lexical

EDUCAÇÃO

Doutorado em Processamento de Linguagem Natural

- Universidade de Sheffield
- 2014-2016

Bacharelado em Ciência da Computação

- Universidade Estadual do Oeste do Paraná
- 2009-2013

Graduação Sanduíche em Ciência da Computação

- Universidade de Sheffield
- 2012

HABILIDADE COM LÍNGUAS

Português:

- Falante nativo

Inglês:

- Fala, leitura e escrita fluentes

HABILIDADE COM LINGUAGENS DE PROGRAMAÇÃO

Avançada:

- Java
- Python

Limitada:

- SQL
- Matlab
- Assembly
- C/C++
- HTML/CSS/Javascript
- NXC (Not eXactly C)
- Pure Data

EXPERIÊNCIA PROFISSIONAL

2016-Atualmente Universidade de Sheffield

Pesquisador Associado em Adaptação de Texto

- Responsável pela liderança do desenvolvimento de novas tecnologias de adaptação de texto no contexto do projeto SIMPATICO, da iniciativa H2020 da União Européia. O objetivo principal do projeto é o desenvolvimento de tecnologias assistivas para diversos públicos alvo, como idosos e falantes não-nativos da língua dominante, no Reino Unido, Itália e Espanha.

06/2015-07/2015 Iconic Translation Machines Ltd.

Pesquisador Desenvolvedor em Estimativa de Qualidade de Tradução

- Responsável pelo desenvolvimento e implantação de métodos inovadores de estimativa de qualidade para a plataforma de tradução automática da empresa Iconic Translation Machines Ltd.

PROJETOS FINANCIADOS DE DESENVOLVIMENTO DE SOFTWARE

2014 ***Incorporando a Ferramenta QuEst à Plataforma Okapi***

- Responsável por liderar o desenvolvimento de um plugin de estimativa de qualidade de traduções para a plataforma Okapi.
- Projeto financiado pela Associação Européia de Tradução de Máquina.

2015 ***Incorporando Recursos a Nível de Palavra à Ferramenta QuEst***

- Responsável por liderar o desenvolvimento dos módulos de extração de características de traduções a nível de palavra para o treinamento de modelos.
- Projeto financiado pela Associação Européia de Tradução de Máquina.

2015 ***Incorporando a Ferramenta IRSTLM à Ferramenta QuEst***

- Responsável por a incorporação da ferramenta IRSTLM de modelos de linguagem na ferramenta para estimativa de qualidade de traduções QuEst.
- Projeto financiado pela empresa bmmt GmbH.

2016 ***Incrementando a Usabilidade da Ferramenta PET***

- Responsável por liderar o desenvolvimento de incrementos de interface e usabilidade para a ferramenta PET (Post-Editing Tool).
- Projeto financiado pela empresa text&form GmbH.

MASSAlign: Alinhamento e Anotação de Documentos Comparáveis

- Uma biblioteca em Python que oferece tecnologias de alinhamento e anotação a nível de palavra de parágrafos e sentenças em documentos comparáveis.
- <https://ghpaetzold.github.io/massalign>

LEXenstein: Uma Plataforma para Simplificação Lexical

- Uma biblioteca em Python que permite a construção e avaliação de diversos tipos de simplificadores lexicais.
- <https://github.com/ghpaetzold/LEXenstein>

QuEst++: Estimativa de Qualidade Multi-Nível

- Ferramenta em Java e Python que facilita o treinamento de modelos de aprendizado de máquina para estimativa de qualidade.
- <https://github.com/ghpaetzold/questplusplus>

EXTLex: Um Analisador Lexical Configurável

- Um analisador lexical em Java para compiladores que permite a detecção de diversos tipos de erros lexicais.
- <http://ghpaetzold.github.io/extlex>

Morph Adorner Toolkit: Adorno Facilitado de Textos

- Ferramenta em Python que oferece uma série de funcionalidades úteis, como inflexão de verbos, substantivos, adjetivos e advérbios.
- <http://ghpaetzold.github.io/MorphAdornerToolkit>

A Survey on Lexical Simplification. JAIR, Volume 60. 2017.

Complex Word Identification: Challenges in Data Annotation and System Performance. Proceedings of the 8th IJCNLP. 2017.

MASSAlign: Alignment and Annotation of Comparable Documents. Proceedings of the 8th IJCNLP. 2017.

Learning How to Simplify From Explicit Labeling of Complex-Simplified Text Pairs. Proceedings of the 8th IJCNLP. 2017.

A Lightweight Regression Method to Infer Psycholinguistic Properties for Brazilian Portuguese. Proceedings of the 20th TSD. 2017.

Lexical Simplification with Neural Ranking. Proceedings of the 15th EACL. 2017.

Understanding the Lexical Simplification Needs of Non-Native Speakers of English. The 26th COLING. 2016

Collecting and Exploring Everyday Language for Predicting Psycholinguistic Properties of Words. The 26th COLING. 2016

Anita: An Intelligent Text Adaptation Tool. The 26th COLING. 2016

Multi-Level Quality Prediction with QuEst++. The 19th EAMT. 2016

SHEF-MIME: Word-level Quality Estimation Using Imitation Learning. The 1st WMT. 2016

SimpleNets: Machine Translation Quality Estimation with Resource-Light Neural Networks. The 1st WMT. 2016

SemEval 2016 Task 11: Complex Word Identification. The 10th SemEval. 2016

SV000gg at SemEval-2016 Task 11: Heavy Gauge Complex Word Identification with System Voting. The 10th SemEval. 2016

PLUMBErr: An Automatic Error Identification Framework for Lexical Simplification. The 1st QATS. 2016

SimpleNets: Evaluating Simplifiers with Resource-Light Neural Networks. The 1st QATS. 2016

Inferring Psycholinguistic Properties of Words. The 15th NAACL. 2016

Benchmarking Lexical Simplification Systems. The 10th LREC. 2016

Unsupervised Lexical Simplification for Non-Native Speakers. The 30th AAAI. 2016

SHEF-NN: Translation Quality Estimation with Neural Networks. The 10th WMT. 2015

LEXenstein: A Framework for Lexical Simplification. The 53rd ACL. 2015

Multi-level Translation Quality Prediction with QuEst++. The 53rd ACL. 2015

Reliable Lexical Simplification for Non-Native Speakers. The 14th NAACL. 2015

EXTLex: An Extensible Lexical Analyser Capable of Detecting Lexical Errors. Revista Eletrônica Científica Inovação e Tecnologia. 2015

Okapi+QuEst: Translation Quality Estimation within Okapi. The 18th EAMT. 2015

Using Positional Suffix Trees to Perform Efficient Tree Kernel Calculation. The 20th NODALIDA. 2015

Text Simplification as Tree Transduction. The 9th STIL. 2013

EXTLex: A Configurable Lexical Analyzer. V Meeting for Computing in Parana. 2013

A Matlab® Remote Control with Voice Command Support for Lego®. V Meeting for Computing in Parana. 2013

Alternative Remote Controls for Lego® Mindstorms® NXT 2.0. IV Meeting for Computing in Parana. 2011

A Remote Control with Bluetooth® Connection for Lego®. IV Meeting for Computing in Parana. 2011

PRÊMIOS

Melhor Formando de Ciência da Computação de 2013

- Concedido pela Universidade Estadual do Oeste do Paraná.

Aluno Destaque em Ciência da Computação de 2013

- Concedido pela Sociedade Brasileira de Computação.

BOLSAS DE ESTUDO CONQUISTADAS

Bolsa do Programa de Educação Tutorial (PET)

- Concedida pelo Ministério da Educação do Brasil.
- 2011-2013

Bolsa de Graduação Sanduíche pelo Ciência Sem Fronteiras

- Concedida pelo Ministério da Educação do Brasil.
- 2012

Bolsa de Doutorado em Ciência da Computação

- Concedida pela Universidade de Sheffield
- 2014-2016