

Veselin Stoyanov

[✉ vesko.st@gmail.com](mailto:vesko.st@gmail.com) | [\(302\) 345-6773](tel:(302)345-6773) | [in LinkedIn](https://www.linkedin.com/in/vesstoyanov) | [📄 Google Scholar](https://scholar.google.com/citations?user=vesstoyanov) | [🌐 vesstoyanov.com](https://www.vesstoyanov.com)

Applied Research Leader excited about turning current powerful AI models into new products that enhance productivity and creativity. Strong track record of innovating in AI and NLP to solve real-world problems. Led teams to create industry-standard pretraining and large LM methods such as RoBERTa, XLM-R and MultiRay and apply them to e.g., reduce the prevalence of hate speech and bullying posts. Experienced in building and motivating high-performing diverse teams and mentoring researchers and engineers.

EXPERIENCE

Tome AI, San Francisco, CA

Head of AI/ML

Apr 2023- Current

Developing new paradigms for AI-powered products in the productivity space.

Facebook / Meta Inc, Menlo Park, CA

Applied Research Scientist Manager

Jul 2018 - Nov 2022

Research Scientist

Jan 2013 - Jul 2018

Project highlights

MultiRay

Service to run multiple very large and accurate models on the same input, and share the majority of the computational costs. MultiRay makes it possible for very accurate self-supervised models to be run on every piece of content. ([paper](#), [blog](#))

Cross-lingual NLP through XLM-R

XLM-R is a state-of-the-art large-scale multilingual language model ([paper](#), [blog](#)). We applied it to extend Integrity classifiers to many languages ([blog](#)). Extended upon previous work on multilingual word embeddings ([blog](#)).

RoBERTa for Integrity

RoBERTa is a state-of-the-art self-supervised method ([paper](#), [blog](#), [blog](#)). Applied it to identifying violations such as hate speech ([blog](#)) and bullying. ([paper](#), [blog](#))

Neural Machine Translation

Shipped the first commercial Neural MT system. Big improvements to quality. ([blog](#), [news](#))

NLP for Search

Shipped several impactful NLP features to Facebook Search including phonetic name search, intent classification and keyword typeahead.

Center for Language and Speech Processing (CLSP), Johns Hopkins University

Assistant Research Scientist

Oct 2010 - Jan 2013

Computing Innovation Fellowship (awarded by CRA) focusing on learning for Structured Prediction.

EDUCATION

Cornell University

Ph.D. in Computer Science

August 2010

Master of Science in Computer Science

May 2005

Advisor: Prof. Claire Cardie. Thesis: *Opinion Summarization: Automatically Creating Useful Representations of Opinions Expressed in Text.*

University of Delaware

Honors BSc, with Distinction in Computer Science

May 2002

Graduated Summa Cum Laude; GPA: 4.00/4.00. Minors in Mathematics and Cognitive Science.

SELECTED PUBLICATIONS

34,000+ citations. Full publication list on [Google Scholar](#) and [Semantic Scholar](#).

[RoBERTa: A Robustly Optimized BERT Pretraining Approach](#)

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, **Veselin Stoyanov**

[BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension](#)

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, **Veselin Stoyanov**, Luke Zettlemoyer

[Unsupervised Cross-lingual Representation Learning at Scale](#)

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, **Veselin Stoyanov**

[XNLI: Evaluating Cross-lingual Sentence Representations](#)

Alexis Conneau, Guillaume Lample, Ruty Rinott, Adina Williams, Samuel R Bowman, Holger Schwenk, **Veselin Stoyanov**

[Supervised Contrastive Learning for Pre-trained Language Model Fine-tuning](#)

Beliz Gunel, Jingfei Du, Alexis Conneau, **Veselin Stoyanov**

[Emerging Cross-lingual Structure in Pretrained Language Models](#)

Alexis Conneau, Shijie Wu, Haoran Li, Luke Zettlemoyer, **Veselin Stoyanov**

[Pretrained Encyclopedia: Weakly supervised knowledge-pretrained language model](#)

Wenhan Xiong, Jingfei Du, William Wang, **Veselin Stoyanov**

[Preserving integrity in online social networks](#)

Alon Halevy, Cristian Canton-Ferrer, Hao Ma, Umut Ozertem, Patrick Pantel, Marzieh Saeidi, Fabrizio Silvestri, **Veselin Stoyanov**

[Empirical risk minimization of graphical model parameters given approximate inference, decoding, and model structure](#)

Veselin Stoyanov, Alexander Ropson, Jason Eisner

[Conundrums in noun phrase coreference resolution: Making sense of the state-of-the-art](#)

Veselin Stoyanov, Nathan Gilbert, Claire Cardie, Ellen Riloff