

Kevin Stowe | PhD.

3158 N. Seminary Ave., 4A, Chicago, IL, 60657 – United States

+1 517 648 8953 • kevincstowe@gmail.com
kevincstowe.github.io

Introduction

I am currently employed as a Research Scientist in Natural Language Processing (NLP) at Educational Testing Services (ETS). I work with the Automated Content Generation team in the field of language generation/generative AI, designing and implementing methods to automatically generate educational content using state-of-the-art NLP techniques including LLMs as well as sequence-to-sequence models. My recent work focuses on fairness and bias issues in educational content generation and automatic prompt-self correction.

My general research interests are in two very different but sometimes overlapping areas: (1) the wide variety of novel, interesting language that people use easily and machines struggle with, and (2) how to ethically and responsibly apply NLP technology to address real-world problems.

Work Experience

Educational Testing Service

Princeton, New Jersey, United States.....

- 2024-Present | Research Scientist (NLP)
- 2022-Present | Associate Research Scientist (NLP)

Automated Content Generation (ACG) team

I develop novel methods for automatically producing content, reducing the time and cost necessary to build high-quality assessments and applications. This is accomplished via both LLM prompting and other deep learning-based language generation to quickly and effectively generate samples used for testing language learners' abilities and providing them with new content for learning: my work combines cutting-edge LLM technology, fine-tuning, prompt engineering, and infrastructure development, as well as rapidly prototyping in an agile environment to deliver new item generation capability.

Responsible AI

I am the responsible AI lead for the ACG team and work to implement bias and fairness reduction for our automated generation processes. This includes running annotation studies, collaborating with domain experts, and building understanding of responsible AI in a rapidly changing environment. I am also a member of the ETS responsible AI research group.

Ubiquitous Knowledge Processing (UKP) Lab

Technical University of Darmstadt, Darmstadt, Germany.....

- 2019-2022 | Postdoc

adv: Prof. Dr. Iryna Gurevych

https://www.informatik.tu-darmstadt.de/ukp/ukp_home

Creative Natural Language Generation

I developed experiments with deep learning methods for language generation, and adding functionality to allow them to produce metaphors, humor, and other challenging types of language. This work involves collaboration with colleagues at TU Darmstadt, as well as thesis students and external partners, and explores a wide breadth of challenges, including manipulating state-of-the-art architectures, collecting appropriate data, and devising evaluation metrics that can handle creative outputs.

Text Analytics for Social Good

I'm involved with a variety of initiatives to use our expertise in NLP to assist with social good projects, ranging from issues of social unrest to public health. This work involves extensive data collection and annotation, and we aim to incorporate citizen science initiatives to alleviate the workload. This collaboration involves computer scientists working on advanced models for data analysis and social scientists providing theory- and task-based guidance.

2013-2019.....

- **Research Associate, CLEAR Lab**
University of Colorado, Boulder, United States

adv: Martha Palmer & Jim Martin

Doctoral Dissertation: Computational Approaches to Metaphor and Syntax

My doctoral dissertation for the University of Colorado is focused on using syntactic features coupled with deep learning to improve automatic metaphor detection. Drawing from evidence that metaphoric meaning can often be derived via syntactic features like argument structure, I've employed information from syntactic parses and lexical resources to improve deep learning models. Through a variety of strategies including improving training data and embedding models, I've improved the performance of state-of-the-art models on multiple metaphor detection tasks.

Computational Language and Education Research Lab (CLEAR)

adv: Martha Palmer

<https://www.colorado.edu/lab/clear/>

I developed and maintained multiple computational lexical resources, including VerbNet and PropBank. My responsibilities included ensuring compatibility with outside resources, implementing new infrastructure, developing interface tools for other researchers, and improving the accuracy, consistency, and coverage of the resources. I worked towards linking VerbNet with the Generative Lexicon, improving consistency among semantic predicates, and improving automatic classification with better annotation.

Communicating Hazard Information in the Modern Environment (CHIME)

adv: Leysia Palen

I served as the natural language processing specialist for the Empowering the Public with Information in Crisis (EPIC) lab at the University of Colorado along with the National Center for Atmospheric Research (NCAR). This project aimed to identify relevant information from social media during natural disasters to assist first responders, government agencies, and affected populations. We developed improved machine learning classification for tweet relevance, as well as building classification based on language and location for predicting evacuation behavior.

Publications

Primary Author.....

- Stowe, Kevin; Longwill, Benny; Franics, Alyssa; Aoyama, Tatsuya; Ghosh, Debanjan; Somasundaran, Swapna. Avoiding Taylor Swift: Identifying Fairness Issues in Automatically Generated Testing Content. Building Educational Applications (BEA) Workshop. 2024. (in submission)
- Stowe, Kevin*; Ghosh, Debanjan*; Zhao, Mengxuan. [Controlled Generation for Language Learning Items](#). Industry track at the 2022 Conference on Empirical Methods for Natural Language Processing (EMNLP). 2023. Abu Dhabi/Online. (*equal contribution)
- Stowe, Kevin; Utama, Prasetya Ajie; Gurevych, Iryna. IMPLI: Exploring Language Models' Performance on Figurative Language. [60th Annual Meeting of the Association for Computational Linguistics \(ACL\)](#). 2022. Online.
- Stowe, Kevin; Beck, Nils; Gurevych, Iryna. [Exploring Metaphoric Paraphrase Generation](#). *25th Annual Conference for Computational Language Learning (CoNLL)*, pg. 323-336. 2021. Online.
- Stowe, Kevin; Chakrabarty, Tuhin; Muresan, Smaranda; Peng, Nanyung; Gurevych, Iryna. [Metaphor Generation with Conceptual Mappings](#). *59th Annual Meeting of the Association for Computational Linguistics (ACL)*, pg.. 6724-6736. 2021. Online.
- Stowe, Kevin; Preciado, Jenette; Conger, Kathryn; Brown, Susan; Kazeminejad, Ghazaleh; Gung, James; Palmer, Martha. [Semlink 2.0: Chasing Lexical Resources](#). *14th International Conference on Computational Semantics (ICWS)*, pg.. 222-227. 2021. Online.
- Stowe, Kevin; Moeller, Sarah; Michaelis, Laura; Palmer, Martha. [Linguistic Analysis Improves Neural Metaphor Detection](#). *23rd Annual Conference for Computational Language Learning (CoNLL)*, pg. 362-371. 2019. New Orleans, Louisiana, US.
- Stowe, Kevin; Palmer, Martha; Anderson, Jennings; Kogan, Marina; Palen, Leysia; Anderson, Kenneth M.; Morss, Rebecca; Demuth, Julie; Lazrus, Heather. [Developing and Evaluating Annotation Procedures for Twitter Data during Hazard Events](#). in *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*, pg. 133-143. 2018. Santa Fe, New Mexico, US.
- Stowe, Kevin; Anderson, Jennings; Palmer, Martha; Palen, Leysia; Anderson, Kenneth M. [Improving Classification of Twitter Behavior During Hurricane Events](#). in *Proceedings of the Workshop on Natural Language Processing for Social Media (SocialNLP)*, pg. 67-75. 2018. Melbourne, Australia
- Stowe, Kevin; Palmer, Martha. [Leveraging Syntactic Constructions for Metaphor Identification and Interpretation](#). in *Proceedings of the Workshop on Figurative Language Processing*, pg. 17-26. 2018. New Orleans, Louisiana, US
- Stowe, Kevin; Paul, Michael J.; Palmer, Martha; Palen, Leysia; Anderson, Kenneth M. [Identifying](#)

and Categorizing Disaster-Related Tweets, in *Proceedings of the International Workshop on Natural Language Processing for Social Media*, pg. 1-6. 2016. Austin, Texas, US

Contributing Author.....

- Ahmed et al., Making Event coreference resolution Tough Again. Metaphorically speaking. 2024. (In submission)
- Klie, Jan-Christoph et al., [Lessons Learned from a Citizen Science Project for Natural Language Processing](#). In Proceedings of the European Association for Computational Linguistics (EACL). 2023. Dubrovnik, Croatia/Online.
- Praminick, Aniket et al., [The challenges of temporal alignment on Twitter during crises](#). In Findings of the Association for Computational Linguistics: EMNLP. 2022. Abu Dhabi, United Arab Emirates.
- Scheunemann, Christoph et al., [Data Collection and Annotation Pipeline for Social Good Projects](#). In Proceedings of the AAAI Fall 2020 AI For Social Good Symposium. 2020. Online.
- Morss, Rebecca et al., [Understanding Weather Forecast Communication, Interpretation, and Use through Analysis of Twitter Data](#). In 29th Conference on Weather Analysis and Forecasting. 2018. Denver, Colorado, US.
- Demuth, Julie L. et al., ["sometimes da #beachlife ain't always da wave": Understanding People's Evolving Hurricane Risk Communication, Risk Assessments, and Responses Using Twitter Narratives](#), In *Weather, Climate, and Society*. 2018.
- Palmer, Martha et al., The Pitfalls of Shortcuts: Tales from the word sense tagging trenches. in *Essays in Lexical Semantics and Computational Lexicography - In Honor of Adam Kilgarriff*. M. Diab, A. Villavicencio, M. Apidianaki, V. Kordoni, A. Korhonen, P. Nakov, M. Stevenson (editors). Springer series Text, Speech and Language Technology. Springer, 2018.
- Morss, Rebecca et al., [Hazardous Weather Predication and Communication in the Modern Information](#). in *Bulletin of the American Meteorological Society*, vol. 98, pg. 2652-2674. 2016.
- Anderson, Jennings et al., [Far Far Away in Far Rockaway: Responses to Risks and Impacts during Hurricane Sandy through First-Person Social Media Narratives](#), in *Proceedings of ISCRAM*, Rio de Janeiro, Brazil, 2016.
- Bonial, Claire et al., [Renewing and Revising SemLink](#), in *The GenLex Workshop on Linked Data in Linguistics*, pg. 9-17, 2013. Pisa, Italy

Dissertation.....

- Stowe, Kevin. [Syntactic and Semantic Improvements to Computational Metaphor Processing](#). Advisors Martha Palmer and Jim Martin. University of Colorado. 2019.

Community

Reviewing.....

- Program Committee, Workshop on Figurative Language Understanding (FigLang) 2024
- Program Committee, Building Educational Applications (BEA) 2024
- Program Committee, Designing Meaning Representations, LREC-COLING 2024
- Program Committee, Industry Track at EMNLP 2023
- Reviewer, Journal of Natural Language Engineering, 2023
- ACL Rolling Review, Nov 2021-Present
- Program Committee, StarSEM 2023
- Reviewer, ACL 2023
- Reviewer, Industry Track at ACL 2023
- Reviewer, Artificial Intelligence Review
- Program Committee, EMNLP 2022
- Program Committee, *SEM 2022
- Program Committee, Figurative Language Workshop, 2022
- Assistant Senior Area Chair (Iryna Gurevych), Area COI Track, Association for Computational Linguistics (ACL), 2021
- Assistant Area Chair (Iryna Gurevych), Sentiment Analysis, Style and Argumentation, Empirical Methods in Natural Language Processing (EMNLP), 2020
- Assistant Area Chair (Iryna Gurevych), Sentiment Analysis, Style and Argumentation, Association for Computational Linguistics (ACL), 2020
- Program Committee, Conference on Computational Natural Language Learning (CoNLL), 2017, Vancouver, Canada
- Program Committee, Corpus Linguistics Fest, 2016, Bloomington, Indiana

Students.....

- Kea Busemann, BS, Evaluating Metaphor Generation
- Elena Leidinger, MS, Exploring Creative Dialogue Agents
- Nils Beck, BS, Controlled Metaphor Generation
- Ruslan Sandler, BS, Natural Language Generation for Humor
- Tobias Blei, BS, Style Transfer for Creative Text Generation

Other.....

- Primary Organizer, Online UKP Retreat, 2020
- Project Manager, Data Analytics Software Project (DASP), TU Darmstadt Summer Semester 2020
- Sponsor, Computer Science Senior Project, Computational Language and Education Research Website Development, 2019
- Student Volunteer, North American Association of Computational Linguistics (NAACL) 2015, Denver, Colorado
- Computer science tutor for Linguistic students, particularly in the aim of improving basic programming skills. Primarily in Python and C++, with other support as required.

Education

- **University of Colorado, Boulder** **Boulder, Colorado**
PhD, Advisors: Martha Palmer, Jim Martin *2013–2019*
Joint degree in Linguistics and Computer Science through the Institute of Cognitive Science (ICS).

Main CS Coursework	Main Ling Coursework
Natural Language Processing	Computational Lexical Semantics
Fundamentals of Programming Languages	Construction Grammar
Machine Learning	Computational Phonology
User-Centered Design	Semantics and Pragmatics
Computation for Research	Phonology, Morphology, Syntax
- **Indiana University** **Bloomington, Indiana**
MA, Linguistics *2009–2011*
Coursework in corpus linguistics, computational approaches to syntax, discrete mathematics and natural language processing.
- **Michigan State University** **East Lansing, Michigan**
BA, Linguistics *2004–2009*
Coursework included core linguistics, advanced semantics and pragmatics, and German and Russian languages.

Technical skills

- **Large Language Models** Experience with prompting, fine-tuning, and evaluation of LLMs. Skilled using the OpenAI interface as well as related tools for fine-grained prompting control.
- **Natural Language Processing** Extensive experience with machine learning and NLP toolkits, including HuggingFace, Spacy, SciKit-Learn, ClearNLP, Stanford Core NLP, Gensim, PyTorch and Keras. Familiar with a wide variety of NLP tasks spanning generation, classification, regression, and unsupervised methods for many domains.
- **Responsible AI/AI for Social Good** Commitment to both implementing and developing research into responsible AI, including ethical use, addressing bias and accessibility issues, and understanding of positive and negative impacts of AI applications, technology, and research.
- **Linguistics** Traditional background in phonetics, phonology, morphology, syntax, and semantics. My primary interests are in lexical semantics, pragmatics, formal logic, and metaphor interpretation.
- **Programming** Proficient in Python and Java, with some experience in C++, Ruby, R, PHP, and SQL. Web development experience with HTML, CSS, Javascript and PHP.
- **Software** Comfortable in Linux, Mac, and Windows environments, with extensive experience with Word/Excel/other office tools and their open-source counterparts. AWS experience including using EC2 instances for model training and inference, S3 for storage, and Sagemaker for deployment. Familiar with FastAPI, Streamlit, and others for prototyping and application experimentation.
- **Management** I've managed a variety of annotation projects, including behavioral annotation of Twitter data and word sense annotation, using Excel as well as custom in-house tools. I've managed coworkers in the generative AI space, and lead special interest groups within academic and industry settings, as well as mentoring junior lab members and supervising BS. and MS. student theses.

- **Professional** My research prioritizes diversity, working with a wide variety of computer scientists, social scientists, and non-academics to accomplish research goals. I'm proficient in research and grant writing, including copy-editing in English. I'm effective at managing time and resources, and developing and executing short- and long-term research plans.

Previous Industry Employment

- **FindMyAudience (<https://findmyaudience.wordpress.com/>)** **Boulder, Colorado**
NLP Consultant *Summer 2015*
I worked for FindMyAudience, a technology startup, to identify possible audiences for authors and publishing companies. Together we developed methods for identifying book similarities and consumer interests from social media and other sources using deep learning and latent semantics models.
- **Avaya Labs** **Westminster, Colorado**
NLP Researcher *Summer 2013*
As a research intern, I did analysis of social media data (Twitter and Facebook) using machine learning algorithms, particularly unsupervised clustering, to determine trends in user interactions with various company partners. I identified differences in positive and negative reactions based on sentiment analysis using topic modelling on social media, allowing for better interaction between companies and their customers.
- **Jackson National Life** **Okemos, Michigan**
Software Developer *2011-2012*
I started as a software trainee and advanced to software developer. I studied Java and SQL to improve software and workflows for company employees.