Seth Ebner

sebner.com
seth.ebner@kensho.com

Updated May 4, 2025

Research Interests

Language understanding (cross-lingual, multilingual, document-level, few-shot, zero-shot), data augmentation, speaker belief, large language models (LLMs), hallucination, factuality, information extraction, computational linguistics

Education

Johns Hopkins University, Baltimore, MD — Aug 2017 – Feb 2025
PhD, Computer Science
Advisor: Benjamin Van Durme
Thesis: *Beliefs About Information Across Languages*

Johns Hopkins University, Baltimore, MD — Aug 2017 – May 2019
MSE, Computer Science
Advisor: Benjamin Van Durme

Washington University in St. Louis, St. Louis, MO — Aug 2013 – May 2017
Bachelor of Science, Computer Science
Bachelor of Science, Electrical Engineering
Cumulative GPA: 4.0/4.0
Major GPA: 4.0/4.0
Valedictorian, Summa Cum Laude – Rank: 1/319

Courses: Natural Language Processing, Linguistic & Sequence Modeling, Pragmatics, Semantics I & II, Syntax, Advanced Topics in Data-Intensive Computing, Parallel Programming

Industry Experience

Kensho Technologies — December 2024 – present
Research Scientist

Google — May 2022 – Aug 2022
Research Intern
Developed methods for using large language models to improve structured extraction from templatic documents in the few-shot setting (TensorFlow).
· Supervisor: James Wendt, Jing Xie, Sandeep Tata

Microsoft — May 2021 – Aug 2021
Research Intern, Text Analytics (*Azure Cognitive Services*)
Developed novel approaches for multilingual document-level sentiment analysis and compared approaches to baselines (PyTorch, AllenNLP). Reported progress in bi-weekly newsletters.
· Supervisor: Benjamin Han

Garmin, Olathe, KS — May 2015 – Aug 2015
Software Engineer Intern
Developed image viewing feature to display images located on SD cards on marine chart plotter multi-function displays (C, GarminOS). Implemented file selection dialog page template to enable developers to create customized menus and to increase code reuse (C, GarminOS).

Pepco Holdings Inc., Newark, DE — Summers 2012 – 2014
Engineering Intern
Forecast power load to predict growth and substation capacity overloads (Cyme, ArcGIS). Developed plan for new capacitor placement as part of smart grid implementation (Cyme, ArcGIS, Microstation).

Publications

[1] Michael Krumdick, Charles Lovering, Varshini Reddy, **Seth Ebner**, and Chris Tanner. No Free Labels: Limitations of LLM-as-a-Judge Without Human Grounding. *arXiv preprint arXiv:2503.05061*, 2025.

[2] Charles Lovering, Michael Krumdick, Viet Dac Lai, **Seth Ebner**, Nilesh Kumar, Varshini Reddy, Rik Koncel-Kedziorski, and Chris Tanner. Language Model Probabilities are Not Calibrated in Numeric Contexts. *arXiv preprint arXiv:2410.16007*, 2024.

[3] Jing Xie, James B Wendt, Yichao Zhou, **Seth Ebner**, and Sandeep Tata. FieldSwap: Data Augmentation for Effective Form-Like Document Extraction. In *2024 IEEE 40th International Conference on Data Engineering (ICDE)*, pages 4722–4732. IEEE, 2024.

[4] Zhengping Jiang, Jingyu Zhang, Nathaniel Weir, **Seth Ebner**, Miriam Wanner, Kate Sanders, Daniel Khashabi, Anqi Liu, and Benjamin Van Durme. Core: Robust Factual Precision Scoring with Informative Sub-Claim Identification. *arXiv preprint arXiv:2407.03572*, 2024.

[5] Miriam Wanner*, **Seth Ebner***, Zhengping Jiang, Mark Dredze, and Benjamin Van Durme. A Closer Look at Claim Decomposition. In *Proceedings of the 13th Joint Conference on Lexical and Computational Semantics (*SEM 2024)*, pages 153–175, Mexico City, Mexico, June 2024. Association for Computational Linguistics.

[6] Shabnam Behzad*, **Seth Ebner***, Marc Marone, Benjamin Van Durme, and Mahsa Yarmohammadi. The Effect of Alignment Correction on Cross-Lingual Annotation Projection. In *Proceedings of the 17th Linguistic Annotation Workshop (LAW-XVII)*, pages 244–251, Toronto, Canada, July 2023. Association for Computational Linguistics.

[7] Mahsa Yarmohammadi, Shijie Wu, Marc Marone, Haoran Xu, **Seth Ebner**, Guanghui Qin, Yunmo Chen, Jialiang Guo, Craig Harman, Kenton Murray, Aaron Steven White, Mark Dredze, and Benjamin Van Durme. Everything Is All It Takes: A Multipronged Strategy for Zero-Shot Cross-Lingual Information Extraction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 1950–1967, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.

[8] Haoran Xu, **Seth Ebner**, Mahsa Yarmohammadi, Aaron Steven White, Benjamin Van Durme, and Kenton Murray. Gradual Fine-Tuning for Low-Resource Domain Adaptation. In *Proceedings of the Second Workshop on Domain Adaptation for NLP*, pages 214–221, Kyiv, Ukraine, April 2021. Association for Computational Linguistics.

[9] Yunmo Chen, Tongfei Chen, **Seth Ebner**, Aaron Steven White, and Benjamin Van Durme. Reading the Manual: Event Extraction as Definition Comprehension. In *Proceedings of the Fourth Workshop on Structured Prediction for NLP*, pages 74–83, Online, November 2020. Association for Computational Linguistics.

[10] **Seth Ebner***, Patrick Xia*, Ryan Culkin, Kyle Rawlins, and Benjamin Van Durme. Multi-Sentence Argument Linking. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8057–8077, Online, July 2020. Association for Computational Linguistics.

[11] **Seth Ebner**, Felicity Wang, and Benjamin Van Durme. Bag-of-Words Transfer: Non-Contextual Techniques for Multi-Task Learning. In *Proceedings of the 2nd Workshop on Deep Learning Approaches for Low-Resource NLP (DeepLo 2019)*, pages 40–46, Hong Kong, China, November 2019. Association for Computational Linguistics.

[12] Arya D McCarthy, Tongfei Chen, and **Seth Ebner**. An exact no free lunch theorem for community detection. In *International Conference on Complex Networks and Their Applications*,

pages 176–187. Springer, 2019.

[13] Yunmo Chen, **Seth Ebner**, Tongfei Chen, Patrick Xia, Elias Stengel-Eskin, Tzu-Ray Su, J. Edward Hu, Nils Holzenberger, Ryan Culkin, Craig Harman, Max Thomas, Thomas Lippincott, Aaron Steven White, Kyle Rawlins, and Benjamin Van Durme. NIST TAC SM-KBP 2019 System Description: JHU/UR Framework. 2019.

[14] Patrick Xia, Elias Stengel-Eskin, Tongfei Chen, **Seth Ebner**, Nils Holzenberger, Ryan Culkin, Pushpendre Rastogi, Xutai Ma, and Benjamin Van Durme. NIST TAC SM-KBP 2018 System Description: JHU/UR Pipeline. 2018.

Research Experience

Speaker Belief
Annotation of speaker belief can enrich the outputs of dialogue, language understanding, and information extraction systems to give a more detailed understanding of speaker intent and reported claims. My work in this area involves designing ontologies, annotation protocols, and models supporting a new perspective on speaker belief.

Large Language Models, Hallucination, Factuality, and Metrology
As LLM-generated text becomes more commonplace, it is increasingly important to evaluate how well-supported such text is by external knowledge sources to mitigate and avoid hallucinations. My work in this area involves identifying and correcting weaknesses in metrics for evaluating textual support and factuality to develop more robust factuality metrics.

Cross-Lingual Information Extraction
Practitioners may be interested in extracting and understanding information from text in a language that there is little or no labeled data for. My work in this area involves designing strategies to make use of data in one language to improve performance in another, often through fine-tuning, data projection, or the use of bitexts.

Multi-Sentence Information Extraction
While almost all of the information extraction work of the past decade focused on single-sentence contexts, complete extraction and understanding requires full document context. My work in this area involves designing multilingual information extraction models that use multi-sentence contexts and creating datasets that support multi-sentence annotations.

Prosody Classification June 2016 – Aug 2016
MIT Lincoln Laboratory
Summer Research Intern, Human Language Technology
Used machine learning to determine feasibility of automatically annotating speech utterances to make text-to-speech output sound more natural (MATLAB). Modeled stress, duration, and pitch of syllables to find correlation among prosodic features.
· Supervisor: Michael Brandstein

Neuromorphic Architecture (Senior design project) Sept 2016 – May 2017
Washington University in St. Louis
Implemented support vector machine based on spiking neuron model for neuromorphic architecture, which mimics biological architecture (C, Raspberry Pi). Wrote program to display real-time spiking patterns (Python).
· Supervisor: Shantanu Chakrabartty

Cache Replacement Policies Sept 2015 – May 2016
Washington University in St. Louis
Investigated performance of deterministic and stochastic criticality-based cache replacement policies. Implemented stochastic replacement policy (Java). Wrote script to simulate load/store instructions of real-time processes for analysis of cache performance (Python).
· Supervisor: Ron Cytron

Binary Integer Programming Sept 2014 – May 2015
Washington University in St. Louis
Investigated effects of adding partial solutions as constraints to initial binary integer programs on execution time of solvers (MATLAB, Octave). Also explored effects of clustering constraints on execution time and output of binary integer program solvers (Ruby).
· Supervisor: Ron Cytron

Service

Journal/Conference Reviewer
· ACL Rolling Review February 2025
· ACL Rolling Review October 2024 (secondary)
· TKDE 2021
· ACL Rolling Review November 2021 (secondary)
· ACL 2019 (secondary)
· ACL 2018 (secondary)
· NAACL 2018 (secondary)
· TACL 2017 (secondary)

JHU Computer Science PhD Admissions Committee 2020 – 2023

North American Computational Linguistics Open Competition Sept 2017 – Sept 2020
Johns Hopkins University
· Organized practice exam sessions for JHU site
· Presented puzzles and discussed solutions
· Proctored competition sessions

Teaching

Event Semantics in Theory and Practice Jan 2021 – May 2021
Johns Hopkins University
Teaching Assistant (Instructors: Benjamin Van Durme, Kyle Rawlins)
· Designed and wrote computational modeling and synthesis paper homework assignments
· Graded homework assignments and weekly quizzes
· Held twice-weekly office hours
· Led help session on using AllenNLP
· Rated "Excellent" by every student in the class

Introduction to Formal Languages and Automata Jan 2017 – May 2017
Washington University in St. Louis
Teaching Assistant (Instructor: Jeremy Buhler)
· Graded homework assignments
· Held weekly office hours

Honors and Awards

Tau Beta Pi
Upsilon Pi Epsilon
IEEE Eta Kappa Nu
Dean's List Fall 2013 – Spring 2017
David H. Levy Outstanding Senior Award Spring 2017
Russell R. Pfeiffer Outstanding Junior Award Spring 2016
Outstanding Sophomore Award Spring 2015
Antoinette Frances Dames Award for Productive Scholarship in Engineering Spring 2015

Activities

Johns Hopkins University Quiz Bowl Aug 2017 – present
Member

Washington University Academic Team Aug 2013 – May 2017
President
· Organized tournaments, twice-weekly practices, and team events
Treasurer
· Oversaw funds, expenses, and reimbursements for tournaments and practices

· Created semesterly budgets for tournaments, practices, and team travel
· Managed invoices for tournaments hosting up to 48 teams

Coder Dojo Sept 2014 – May 2016
Volunteer
· Taught middle school students web development

Skills

Languages: Python, Java, C, C++, MATLAB, Ruby, VHDL
Frameworks + Tools: PyTorch, TensorFlow, HuggingFace, Mechanical Turk, Jupyter, Colab