Chris Kedzie

NLP · ML · Generative AI Researcher

in LinkedIn · ○ Github · 8 Google Scholar

Brooklyn, New York

christopher.r.kedzie@gmail.com

+1 (925) 323-1837

EXPERIENCE

Microsoft, Inc.   *June 2022 – present*
Principal Researcher, MSAI   *Sept. 2023 – present*

- **Science lead for enterprise Copilot metrics team.** Conduct experiments to ensure the quality metrics powered by large language models (LLM) continue to function effectively amidst platform changes. Provide guidance on offline A/B experiments. Communicate metrics issues to leadership and Copilot feature team stakeholders. Consult on human judge/LLM metrics calibration workstreams.

- **Directly responsible individual (DRI) for enterprise Copilot online quality metrics.** Design first enterprise Copilot quality metrics for online A/B testing. Implement AML pipelines for LLM-powered extraction of user satisfaction signals from Copilot conversations. Design distillation pipelines to scale LLM evaluation to all worldwide traffic. Work with metrics architect and Bing Copilot metrics teams to bring consistent evaluation suite across all Copilots.

- **Research intern mentor.** Mentor two Ph.D. interns on topics of LLM model probing for hallucination detection and aligning human judges to LLM-based evaluators with papers to appear at ACL 2024.

Senior Researcher, Semantic Machines   *June 2022 – Sept. 2023*

- **Science lead for enterprise Copilot metrics team.** Developed novel task completion based evaluation metrics as well as adapted Bing Chat response quality metrics. Designed metrics guidelines for offline A/B testing framework to support contributing feature teams while ensuring new features did not regress product quality. Led the design of LLM powered unit testing for enterprise Copilot. Designed LLM orchestrator tool invocation metrics.

- **Project member of fast prototyping team for speculative semantic parsing and conversational AI projects.** Prototyped conversational assistant that could translate natural language utterances to Microsoft's internal graph API language using LLMs with grammar constrained decoding. Implemented M365 content retrieval and evaluation framework for GPT-4 powered assistant.

- **Science lead for simulated user project.** Led synthetic data generation project to create additional training data for semantic parsing models. Implemented experiments using in-context prompting for data generation and subsequently trained semantic parsing models in both PyTorch and internal Scala-based ML framework. Worked with data and engineering team to build front-end prototype to support ad-hoc data generation for new parser functionality.

Rasa Technologies, Inc.   *Jan. 2021 – June 2022*
Senior Machine Learning Researcher   *Jan. 2022 – June 2022*

- **Contributor to the Rasa framework**, an open-source python-based ML library for implementing conversational AI assistants (18.3k stars on GitHub and over 50 million downloads). Worked closely with product teams to explore new machine learning features. Collaborated on the implementation of the UnexpecTED ML model to detect unusual dialogue turns, aiding in automatic problem identification for debugging conversational assistants.

- **Lead researcher of user simulation framework for system testing.** Led the development of an internal user simulation framework utilizing LLMs to generate user responses based on descriptions of user goals and/or persona. Facilitated rapid evaluation of novel dialogue system features across research and engineering teams through the user simulation framework, minimizing the need for extensive human evaluation.

Machine Learning Researcher   *Jan. 2021 – Dec. 2022*

- **Contributor to the Rasa framework.** Designed and implemented a template-based response generation system on top of an ML powered conversation state manager, that enabled developers to flexibly delegate response generation to an ML dialogue manager or override with custom business logic. Wrote documentation and engaged in developer relations (YouTube videos and answering forum questions) to help introduce the feature with the developer community.

Facebook Research   *Summer 2017*
Research Intern

- **Applied Machine Learning group intern for clickbait and hate speech detection models.** Developed an adversarial training method to enhance the performance of convolutional neural network text classifiers. Implemented the adversarial training architecture using PyTorch and Caffe2 frameworks. Integrated the developed method into the internal FBLearner system to facilitate experiment tracking and execution, and for reuse by other teams.

Microsoft Research   *Summer 2015*
Research Intern

- **Research intern on scalable summarization systems for streaming news project.** Designed and implemented a text summarization system to deliver concise news updates in real-time using imitation learning. The system achieved top performance in the Temporal Summarization Track of the 2015 Text Retrieval Conference and I was invited to present our model.

SKILLS

Python, PyTorch, Tensorflow, SciKit-Learn, Azure ML, Spark/PySpark, SQL, Streamlit, JavaScript

EDUCATION

Columbia University   New York, NY
Ph.D., Computer Science.   *2014 – 2020*
    Dissertation: Salience Estimation and Faithful Generation: Modeling Methods for
    Text Summarization and Generation. 2021.
M.S., Computer Science.   *2012 – 2014*

Loyola Marymount University   Los Angeles, CA
B.A., Music & Recording Arts Double Major.   *2004 – 2008*

AWARDS

2021	Columbia U. Comp. Sci. Dept. Michelman Award (community service).
2020	EMNLP. Outstanding Reviewer.
2019	INLG. Best Paper Award.
2019	NYAS Natural Language, Dialog and Speech Symposium. STAR Talk 2nd Place Prize.

Evaluation and Interpretability

- LLM-Rubric: A Multidimensional, Calibrated Approach to Automated Evaluation of Natural Language Texts. ACL 2024. Helia Hashemi, Jason Eisner, Corby Rosset, Benjamin Van Durme, and **Chris Kedzie**.

- Do Androids Know They're Only Dreaming of Electric Sheep? ACL Findings 2024. Sky CH-Wang, Benjamin Van Durme, Jason Eisner, and **Chris Kedzie**.

- Content Selection in Deep Learning Models of Summarization. EMNLP 2018. **Chris Kedzie**, Kathleen McKeown, and Hal Daumé III.

Text Generation

- Lightweight Decoding Strategies for Increasing Specificity. Preprint 2021. Katy Gero, **Chris Kedzie**, Savvas Petridis, and Lydia B. Chilton.

- Controllable Meaning Representation to Text Generation: Linearization and Data Augmentation Strategies. EMNLP 2020. **Chris Kedzie** and Kathleen McKeown.

- A Good Sample is Hard to Find: Noise Injection Sampling and Self-Training for Neural Language Generation Models. INLG 2019. **Chris Kedzie** and Kathleen McKeown.

- Low-Level Linguistic Controls for Style Transfer and Content Preservation. INLG 2019. Katy Gero, **Chris Kedzie**, Jonathan Reeve, and Lydia Chilton.

Text Summarization

- Cross-language Sentence Selection via Data Augmentation and Rationale Training. ACL 2021. Yanda Chen, **Chris Kedzie**, Suraj Nair, Petra Galuscakova, Rui Zhang, Douglas Oard, and Kathleen McKeown.

- Real-Time Web Scale Event Summarization Using Sequential Decision Making. IJCAI 2016. **Chris Kedzie**, Fernando Diaz, and Kathleen McKeown.

- Predicting Salient Updates for Disaster Summarization. ACL 2015. **Chris Kedzie**, Kathleen McKeown, and Fernando Diaz.

NLP for Social Science

- Multimodal social media analysis for gang violence prevention. ICWSM 2019. Philipp Blandfort, Desmond U. Patton, William R. Frey, Svebor Karaman, Surabhi Bhargava, Fei-Tzin Lee, Siddharth Varia, **Chris Kedzie**, Michael B. Gaskell, Rossano Schifanella, Kathleen McKeown, Shih-Fu Chang.

- Detecting gang-involved escalation on social media using context. EMNLP 2018. Serina Chang, Ruiqi Zhong, Ethan Adams, Fei-Tzin Lee, Siddharth Varia, Desmond U. Patton, William R. Frey, **Chris Kedzie**, Kathleen McKeown.

Speech and Translation

- Segmenting Subtitles for Correcting ASR Segmentation Errors. EACL 2021. David Wan, **Chris Kedzie**, Faisal Ladhak, Marine Carpuat, Kathleen McKeown.

- Incorporating Terminology Constraints in Automatic Post-Editing. WMT 2020. David Wan, **Chris Kedzie**, Faisal Ladhak, Elsbeth Turcan, Petra Galuscakova, Elena Zotkina, Zhengping Jiang, Peter Bell and Kathleen McKeown.