

KANISHKA MISRA

Email: kanishka@ttic.edu **Website:** <https://kanishka.website/>

Last Updated: September 29, 2024

Academic Positions

2024–current	Toyota Technological Institute at Chicago Research Assistant Professor <i>Endowed research faculty position for up to 3 years</i>
2023–2024	The University of Texas at Austin Postdoctoral Research Fellow, Department of Linguistics Advisor: Dr. Kyle Mahowald

Education

Purdue University, West Lafayette, USA

Ph.D. in Computer Information Technology, 2018–2023

Dissertation: *On Semantic Cognition, Inductive Generalization, and Language Models*

Advisor: Dr. Julia Taylor Rayz

Committee: Dr. Allyson Ettinger, Dr. Victor Raskin, Dr. Jin Wei Kocsis, Dr. John Springer

Purdue University, West Lafayette, USA

M.S. in Computer Information Technology, 2020

Thesis: *Exploring Lexical Sensitivities in Word Prediction Models: A case study on BERT*

Advisor: Dr. Julia Taylor Rayz

Note: Work performed alongside requirements for Ph.D.

Purdue University, West Lafayette, USA

B.S. *with distinction*. Computer Information Technology, 2014–2018

Minor in Statistics

Fellowships and Assistantships

2023	Bilsland Fellowship. Dissertation Fellowship awarded by Purdue University.
2022–2023	Purdue Graduate Student Mentoring Fellowship. Selected to understand and improve the advising relationship between faculty and students at Purdue University. Award: \$5,000 in research and travel funds.
2021–2022	Research Assistantship funded through NSF EAGER Grant number 2039605. Title: <i>AI-based Humor-integrated Social Engineering Training</i> . Contribution: Co-wrote the “Technical Contribution” section, and served as key personnel. PI: Julia Taylor Rayz, Co-PI: Ida B. Ngambeki
2018–2019	Purdue Research Foundation (PRF) Fellowship. Title: <i>Computational Analysis of Online Predatory Texts</i> . Contribution: Wrote the proposal in its entirety. Mentor: Julia Taylor Rayz.

Industry Experience

- Fall 2022 **Google Research - Research Intern**
Project: Triggering Multi-Hop Reasoning in LLMs with Soft-prompts.
Host(s): Siamak Shakeri and Cicero Nogueira dos Santos.
- Summer 2021 **Pythonic AI - NLP Engineering/Research Intern**
Project: Integrating Biomedical Commonsense into Language Models.
Host: Baoqiang Cao, CTO and Co-founder.

Work In Progress and Preprints

- 2024 Juan Diego Rodriguez, Aaron Mueller, **Kanishka Misra**. Characterizing Category-based Inferences in Language Models. *In preparation*.
- 2024 **Kanishka Misra** and Najoung Kim. [Generating novel experimental hypotheses from language models: A case study on cross-dative generalization](#). *Under Review*.
- 2022 **Kanishka Misra**. [minicons: Enabling Flexible Behavioral and Representational Analyses of Transformer Language Models](#). Demo Paper. *Preprint/Demo*.

Peer-reviewed Publications

- 2024 **Kanishka Misra**, Allyson Ettinger, and Kyle Mahowald. [Experimental Contexts Can Facilitate Robust Semantic Property Inference in Language Models, but Inconsistently](#). *EMNLP 2024*.
- 2024 **Kanishka Misra** and Kyle Mahowald. [Language Models Learn Rare Phenomena from Less Rare Phenomena: The Case of the Missing AANNs](#). *EMNLP 2024*.
- 2023 **Kanishka Misra**, Cicero Nogueira dos Santos, Siamak Shakeri. [Triggering Multi-hop reasoning in language models using soft-prompts and random walks](#). *Findings of ACL 2023*
- 2023 Koustuv Sinha*, Jon Gauthier*, Aaron Mueller[†], **Kanishka Misra**[†], Keren Fuentes, Roger Levy, Adina Williams. [Language model acceptability judgments are not always robust to context](#). *ACL 2023*. **Recipient of Outstanding Paper Award**. *,[†]: Equal Contribution
- 2023 Freda Shi*, Xinyun Chen*, **Kanishka Misra**, David Dohan, Ed Chi, Nathan Scharli, Denny Zhou. [Large Language Models Can Be Easily Distracted by Irrelevant Context](#). *ICML 2023*. *: Equal Contribution
- 2023 **Kanishka Misra**, Julia Taylor Rayz, Allyson Ettinger. [COMPS: Conceptual Minimal Pair Sentences for Testing Robust Property Knowledge and its Inheritance in Pre-trained Language Models](#). *EACL 2023*. **Recipient of Best Paper Award**.
- 2022 **Kanishka Misra**, Julia Taylor Rayz, Allyson Ettinger. [A Property Induction Framework for Neural Language Models](#). *44th Annual Conference of the Cognitive Science Society*.
- 2022 **Kanishka Misra**, Julia Taylor Rayz. [LMs Go Phishing: Adapting Pre-trained Language Models to Detect Phishing Emails](#). *IEEE/ACM Web Intelligence Conference*.

- 2022 **Kanishka Misra**. [On Semantic Cognition, Inductive Generalization, and Language Models](#). AAAI 2022 Doctoral Consortium, Vancouver, Canada.
- 2021 **Kanishka Misra**, Allyson Ettinger, Julia Taylor Rayz. [Do Language Models learn typicality judgments from text?](#) 43rd Annual Conference of the Cognitive Science Society.
- 2021 **Kanishka Misra**, Julia Taylor Rayz. [Finding fuzziness in Neural Network models of Language Processing](#). Annual Meeting of the North American Fuzzy Information Processing Society 2021. **Honorable Mention for Best Student Paper**.
- 2020 **Kanishka Misra**, Allyson Ettinger, Julia Taylor Rayz. [Exploring BERT’s Sensitivity to Lexical Cues using Tests from Semantic Priming](#). Findings of the Association for Computational Linguistics: EMNLP 2020.
- 2020 Qingyuan Hu, Yi Zhang, **Kanishka Misra**, Julia Taylor Rayz. [Exploring Lexical Irregularities in Hypothesis-Only Models of Natural Language Inference](#). 2020 IEEE 19th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC).
- 2020 **Kanishka Misra**, Julia Taylor Rayz. [An Approximate Perspective on Word Prediction in Context: Ontological Semantics meets BERT](#). Annual meeting of the North American Fuzzy Information Processing Society 2020. Online.
- 2019 **Kanishka Misra**, Hemanth Devarapalli, Tatiana Ringenberg, Julia Taylor Rayz. [Authorship Analysis of Online Predatory Conversations using Character Level Convolution Neural Networks](#). 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy.
- 2019 Tatiana Ringenberg, **Kanishka Misra**, Julia Taylor Rayz. [Not So Cute but Fuzzy: Estimating Risk of Sexual Predation in Online Conversations](#). 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)., Bari, Italy. **(joint first author)**
- 2019 Tatiana Ringenberg, **Kanishka Misra**, Kathryn C. Seigfried-Spellar, Julia Taylor Rayz. [Exploring Automatic Identification of Fantasy-Driven and Contact-Driven Sexual Solicitors](#). 2019 Third IEEE International Conference on Robotic Computing (IRC)., Naples, Italy.
- 2019 Kathryn C. Seigfried-Spellar, Marcus K Rogers, Julia T Rayz, Shih-Feng Yang, **Kanishka Misra**, Tatiana Ringenberg. [Chat analysis triage tool: Differentiating contact-driven vs. fantasy-driven child sex offenders](#). Forensic Science International, 2019.

Peer-reviewed Abstracts

- 2023 **Kanishka Misra** and Najoung Kim. [Abstraction via exemplars? A representational case study on lexical category inference in BERT](#). BUCLD 48
- 2020 **Kanishka Misra**, Allyson Ettinger, Julia Taylor Rayz. [Exploring BERT’s lexical relations using Semantic Priming](#). CogSci 2020 [poster]
- 2019 **Kanishka Misra**, Hemanth Devarapalli, Julia Taylor Rayz. [L1 Influence on Content Word errors in Learner English Corpora: Insights from Distributed Representation of Words](#). CogSci 2019, Montréal, Canada. [poster]

Presentations and Invited Talks

Fall 2024	<i>Generalization from Indirect Evidence in Language Models</i> Guest lecture for UChicago course: <i>Do LLMs Generalize?</i> (taught by Ari Holtzman) Research @ TTIC Series.
Summer 2024	<i>Controlled Rearing of Language Models can reveal Linguistic Insight</i> (Invited Talk) Princeton Psychology of Language Lab (Adele Goldberg's group)
Spring 2024	<i>Controlled Rearing of Language Models can reveal Linguistic Insight</i> (Invited Talk) SynSem Seminar, UT Austin Linguistics. South by Semantics Workshop, Austin
Fall 2023	<i>Analyzing Robust Conceptual Knowledge and Property Inheritance in Language Models</i> (Invited Talk) Guest lecture for <i>Topics in Natural Language Processing</i> (taught by Eunsol Choi) Toyota Technological Institute at Chicago - Research Series. C.Psyd Lab (Marten van Schijndel's lab), Cornell University.
Fall 2023	<i>Language model acceptability judgments are not always robust to context</i> UT Austin NLL Reading Group.
Spring 2023	<i>COMPS: Conceptual Minimal Pair Sentences for testing Robust Property Knowledge and its Inheritance in Pre-trained Language Models</i> EACL 2023, Main Conference
Spring 2023	<i>On the synthetic semantic cognition of language models</i> (Invited Talk) EvLab and LINGO lab. MIT. Toyota Technological Institute at Chicago.
Fall 2022	<i>Conceptual Minimal Pairs for testing Robust Property Knowledge and its Inheritance in Pre-trained Language Models</i> (Invited Talk) CompLing Lab, University of Chicago Human and Machine Learning Lab. New York University Computation and Psycholinguistics Lab. New York University CompLang/Ev Lab joint meeting. MIT Language Evolution, Acquisition and Processing Workshop. University of Chicago
Fall 2022	<i>Triggering Multi-Hop Reasoning in LLMs using Soft-Prompts</i> Prompt-tuning sync. Google Research.
Spring 2022	<i>On Semantic Cognition, Inductive Generalization, and Language Models</i> AAAI 2022 Doctoral Consortium. Vancouver (held online).
Fall 2021	<i>Finding fuzziness in Neural Network Models of Language Processing</i> NAFIPS 2021 (held online).
Summer 2021	<i>Do Language Models Learn Typicality Judgments from Text?</i> CogSci 2021 Main Conference (held online).

Honors and Awards

- 2024 **Google Cloud Credits. Co-PI along with Najoung Kim.** *Google.* **Amount:** \$5000.
- 2023 **Paula Menyuk Travel Award.** *BUCLD 48.* **Amount:** \$385.
- 2023 **Outstanding Paper Award.** *ACL 2023.*
- 2023 **Best Paper Award.** *EACL 2023.* **Amount:** \$1000.
- 2023 **Bilsland Fellowship.** *Purdue University.* **Amount:** \$25,000 + Tuition (1 year).
- 2022 **Best Student Poster (runner-up) PPI Holistic Safety and Security Research Impact area.** **Amount:** \$250.
- 2022 **Fellow,** *Purdue Graduate Student Mentoring Fellows Program.* **Amount:** \$5,000 in research funds.
- 2021 **Honorable Mention for Best Student Paper,** *North American Fuzzy Information Processing Society.* **Amount:** \$100.
- 2019 **Holistic Safety and Security Research Travel Grant,** *Purdue Polytechnic Institute.* **Amount:** \$500.
- 2019 **CIT Research Travel Grant Award,** *Purdue CIT.* **Amount:** \$1200 (CogSci 2019), \$600 (IEEE-SMC 2019).
- 2019 **Best HSS Poster Presentation,** *CERIAS Symposium.* Award presented by committee on Holistic Safety and Security (HSS) research impact area. [[link](#)].
- 2019 **Conference Travel Award,** *Chicago R Unconference.* **Amount:** \$150.
- 2018 **PRF Fellowship,** *Purdue Research Foundation.* Covered two semesters worth of graduate school, in addition to stipend.
- 2018 **Best Poster Award - PPI,** *Purdue Office of Undergraduate Research Expo.* **Amount:** \$250. [[link](#)]
- 2018 **Research Scholarship,** *Purdue Office of Undergraduate Research.* **Amount:** \$500.
- 2017 **First Place.** *Indy Civic Hackathon.* **Amount:** \$2000 split across 4 team members.

Teaching

Guest Lecturer - *Natural Language Technologies* (CNIT 519)

Timeline: Fall 2019, Fall 2020, Spring 2022

Course Professor: Dr. Julia Taylor Rayz

- Two lectures on Neural Network models of Natural Language Processing
- Developed two assignments on language model interpretability and evaluation.

Teaching Assistant - *Database Fundamentals* (CNIT 272)

Timeline: Fall 2019, Spring 2020, Fall 2020

Course Professor: Dr. Dawn D. Laux

Developed lecture videos and taught fundamentals of relational databases and SQL to three lab sections (≈ 70 students on average across three semesters).

Instructor Rating: 4.8 (on average across three semesters)

Mentorship

- 2024–** **Lalchand Pandia** (PhD Student at the University of Chicago). **Topic(s):** *Investigating distributional signals of hypernymy in Language Models.*
- 2023–** **William Sheffield** (PhD Student at UT Austin). **Topic(s):** *Discourse-sensitive properties of Language Models*
- 2023–** **Juan Diego Rodriguez** (PhD Student at UT Austin). **Topic(s):** *Conceptual Organization and Inference in Language Models*
- 2023–** **Sriram Padmanabhan** (Undergraduate at UT Austin). **Topic(s):** *Sampling sensitivity in Language Models*
- 2023–24** **Jwalanthi Ranganathan** (Undergraduate at UT Austin). **Topic:** *Projecting Contextualized Word Embeddings onto Interpretable Semantic Spaces.*
- 2023–** **Daniel Brubaker** (Undergraduate at UT Austin). **Topic:** *Exploring pragmatic inference in LMs with Discourse Connectives.*
- 2020** **Qingyuan “Carol” Hu and Yi Zhang** (Undergraduates). **Topic:** *Exploring Lexical Irregularities in Hypothesis-only Models of Natural Language Inference.* **Outcome:** Publication in *IEEE ICC* CI 2020*, and a presentation at *PURC 2020*, which was awarded second place across all students from the Purdue Polytechnic Institute.
- 2018–19** **John Phan** (Undergraduate). **Topic:** *Gender Bias in Word Embeddings.* Awarded NSF REU scholarship. **Outcome:** Two poster presentations.

Reviewing

- Grants** National Science Foundation - SBE (2024)
- Journal** *Journal of Memory and Language* (2024) $\times 2$; *Computational Linguistics* (2024); *PNAS Nexus* (2024); *Cognitive Science* (2023); *Open Mind* (2023)

Conference	CogSci (2024); EMNLP (2023); CoNLL (2023); ACL (2023); EACL (2023); CogSci (2020, 2021, 2022); CoNLL (2021, 2022); ARR (2021, 2022); *SEM 2022; EMNLP 2020; IJCAI 2020; *SEM 2019; IEEE-IRC 2019.
Workshop	Blackbox NLP (2023); ICML Workshop on Large Language Models and Cognition (2024)
Book	Chapman & Hall/CRC Press Statistics Series (2020, 2021).

Service

- **Co-organizer**, UChicago and TTIC Joint NLP Seminar Series (2024–).
- **Area Chair/Meta Reviewer**, ACL Rolling Review (ARR) 2024 (Feb, April, June).
- **Area Chair**, *Linguistic Theory, Cognitive Modeling, and Psycholinguistics*. LREC-COLING 2024.
- **Session Chair**, *Linguistic Theory, Cognitive Modeling, and Psycholinguistics*. ACL 2023 (on-site).
- **Organizer**, *Neural Nets for Cognition*. Discussion group at CogSci 2022.
- **Local Arrangements Chair**, *Annual Meeting of the North American Fuzzy Information Processing Society 2021 (NAFIPS 2021) held at Purdue University*.
- **Program Committee**: CoNLL (2021, 2022).
- **Volunteer**, *36th AAAI Conference on Artificial Intelligence*.
- **Graduate Student Advisor**, *Purdue CIT Student Council*.
- **Organizer**, *Undergraduate Research Panel*, Purdue CIT.

Skills

Programming	Python, R, pytorch, jax, SQL, \LaTeX
Natural Languages	English (near-native), Hindi (near-native), Gujarati, Odiya

Software Developed

minicons A toolkit to facilitate behavioral and representational analyses of transformer-based language processing models. [[github](#)]

Professional Affiliations

- Association of Computational Linguistics (ACL)
- Cognitive Science Society (CogSci)
- Society for Mathematical Psychology (MathPsych)