Dr. Damir Cavar

Associate Professor
Indiana University at Bloomington

Phone: (734) 709-6039
Email: dcavar@iu.edu
Homepage: http://damir.cavar.me/
NLP Lab page: https://nlp-lab.org/
GitHub page: `https://github.com/dcavar`
ORCID: `orcid.org/0000-0002-1262-5927`

Curriculum Vitae — 2024–2025

Current Positions and Roles

- **Associate Professor** at Indiana University Bloomington, tenure home Department of Linguistics (former founder and director of the Computational Linguistics Program at IU, back then a joint Cognitive Science, Computer Science, and Mathematics program).
- Since 2024, **Fellow of the Center for Quantum Technologies** (CQT), an NSF Industry/University Cooperative Research Center (IUCRC), including Indiana University, Purdue University, and the University of Notre Dame.
- Since 2024, **Member of the Quantum Science and Engineering Center** (QSEc) at Indiana University.
- Since 2023, **Core AI Faculty at the Luddy Artificial Intelligence Center** at Indiana University.
- Since 2018, **Fellow at the Center for Applied Cybersecurity Research (CACR)** at Indiana University.
- Since 2018, **Data Science Faculty in the Luddy School**, Computer Science Department, Indiana University.
- **Faculty Member (Voting Member)** in the Cognitive Science Program, College of Arts and Sciences (COAS), Indiana University.

Previous Positions and Roles

- 2013–2014 – Director of the Institute for Language Information and Technologies (ILIT), at Eastern Michigan University
- 2018 – Cofounder of startup Semiring Inc. (USA); 2007–2009 Cofounder of MORE IT (Croatia); Cofounder of L3C (Germany)
- 2006–2010 Professor positions at the University of Konstanz (Germany), University of Zadar (Croatia), University of Nova Gorica (Slovenia)
- 2003–2006 Assistant Professor, Indiana University, Director of Computational Linguistics

Education

- **2001: PhD, University of Potsdam, Germany, *magna cum laude***

Affiliation

- Luddy Artificial Intelligence Center Core Faculty, Indiana University Bloomington
- IU Quantum Science and Engineering Center (QSEc)
- Data Science Faculty in the Luddy School of Informatics, Computing, and Engineering
- Center for Applied Cybersecurity Research (CACR) Fellow at Indiana University Bloomington
- Voting Member of the Cognitive Science Program at Indiana University Bloomington

Membership

- Association for Computing Machinery (ACM), Senior Lifetime Member
  - Special Interest Group for Artificial Intelligence (SIGAI)
  - Special Interest Group on Management of Data (SIGMOD)
- Institute of Electrical and Electronics Engineers (IEEE) (Senior Member status in process)
  - IEEE Computer Society Member
  - IEEE Systems Council
- Association for Computational Linguistics (ACL)
- The Association for Logic, Language and Information (FoLLI)

Most Recent Grant Activities 2024–2025

Grants 2024–2025 funded

- July 2024 – May 2025, Center for Quantum Technologies (CQT) grant for *Quantum-Natural Language Processing (NLP) and Machine Learning (ML)*, this is an NSF Center jointly managed by Indiana University, Purdue University, and the University of Notre Dame, full funding for one PhD-student and travel.

Grants 2024–2025 submitted

- 2024, NSF NRT grant for developing and expanding a Quantum Information Science and Engineering (QISE) program at Indiana University. Applied as a multi-disciplinary program that involves the sciences in the College of Arts and Sciences, and the departments of Intelligent Systems Engineering and Computer Science at the Luddy School of Informatics, Computing, and Engineering.
- 2025, NIH R01 grant together with colleagues from the School of Public Health at Indiana University Bloomington on Social Network Studies of drug-related conversations and themes on social media using AI and NLP technologies.
- 2025, Belgian Science Policy Office (BELSPO) grant in cooperation with the National Institute for Criminalistics and Criminology in Brussels, Belgium on applying Large Language Models, Knowledge Graphs, and Ontologies in Criminological Research. Cooperative project between IU and various universities in Canada and Belgium. This is a non-funded research opportunity to prepare a grant proposal, potentially together with colleagues from the IU Department of Criminology and Criminal Justice.
- Co-PI: Humanities and Artificial Intelligence Virtual Institute (HAVI) - Schmidt Sciences, joint proposal with a colleague from Cognitive Science at Indiana University at Bloomington and Computer Science at the Luddy School of Informatics, Computing, and Engineering. Submitted in Spring 2025.

Teaching

2025

- CSCI-B 659 Advanced Machine Learning Techniques / Topics in Artificial Intelligence
- CSCI-B 659 Advanced Natural Language Processing / Topics in Artificial Intelligence
- CSCI-B 659 Seminar on Knowledge Graphs, Large Language Models, Graph-based Reasoning, Retrieval Augmented Generation (RAG), and Ontologies / Topics in Artificial Intelligence

2024

- CSCI-B 659: Applying Machine Learning Techniques in Computational Linguistics / Topics in Artificial Intelligence
- CSCI-B 659 Advanced Natural Language Processing / Topics in Artificial Intelligence
- Generative AI and Symbolic Knowledge Representations: Large Language Models, Knowledge, and Reasoning (July/August 2024) Course at the 35th European Summer School in Logic, Language and Information ESSLLI 2024 in Leuven, Belgium.
- CSCI-B 659 Applying Machine Learning Techniques in Computational Linguistics / Topics in Artificial Intelligence
- HON-H 240 (Hutton Honors College) Language, Intelligence, and the Machine, Undergraduate introduction to Artificial Intelligence

Related activities

- Organizer of the Quantum AI and NLP 2025 conference, an ACM in-cooperation conference, with post-proceedings publishing with Springer; raised more than $ 25,000 for the conference, invited Vice Presidents of IBM and Microsoft, the academic director of NVIDIA CUDA Quantum, and the scientific director of Quantinuum, among others.
- Directing the Natural Language Processing (NLP) Lab at IU for more that 6 years: `https://nlp-lab.org/`.
- Directing the Quantum AI and NLP Study Group at IU since 2023 `https://nlp-lab.org/quantumnlp/`.
- Industry Partnership Activities (Indiana University and Cook Medical), initial meetings and preparation of joint projects.
- I received a Client-based International Projects (CLIP) Program grant for Fall 2025 to include the students of a seminar in international projects. The projects involved cooperation with a medical team in India working on medical AI technologies, a criminology and criminal justice team in the European Union (Brussels) working on AI and NLP for forensic analysis, and an advertising team in Croatia working on automatic ad classification.

Publications 2024–2025

1. Damir Cavar, Shane A. Sparks, James Bryan Graves, Koushik Reddy Parukola, Billy G. Dickson (to appear in 2025) The Natural Language Qu Kit – NLQK for Quantum Natural Language Processing and AI. Submitted for review. (See NLQK)

2. Damir Cavar, Koushik Reddy Parukola, James Bryan Graves, Shane A. Sparks (accepted for publication, 2025) Old Wine in New Bottles: Using Classical Word Embeddings in Gate-Based Quantum NLP Systems.

3. Damir Cavar, James Bryan Graves, Shane A. Sparks (submitted for review, 2025) On Quantum Encoding of Dense Vector Embeddings for AI. Submitted for review.

4. Damir Cavar, Koushik Reddy Parukola (2025) *Word and Text Similarity Using Classical Word Embeddings in Quantum NLP Systems.* Satellite Workshop: Quantum Machine Learning in Signal Processing and Artificial Intelligence at the 2025 IEEE International Conference on Acoustics, Speech, and Signal Processing. Hyderabad, India.

5. Muhammad S. Abdo, Yash Hatekar, and Damir Cavar (2025) *AMWAL: Named Entity Recognition for Arabic Financial News.* In Proceedings of the FinNLP-FNP-LLMFinLegal 2025 workshop at COLING 2025.

6. Damir Cavar and Chi Zhang (2024) *Semantic Similarities using Classical Embeddings in Quantum NLP.* In Proceedings of the *IEEE Quantum Week 2024*, Montreal, Canada, September 2024.

7. Chi Zhang, Akriti Kumari, Damir Cavar (2024) *Entangled Meanings: Classification and Ambiguity Resolution in Near–Term QNLP.* In Proceedings of the *IEEE Quantum Week 2024*, Montreal, Canada, September 2024.

8. Damir Cavar, Zoran Tiganj, Ludovic Mompelat, Billy Dickson (2024) "Computing Ellipsis Constructions: Comparing Classical NLP and LLM Approaches." *Society for Computation in Linguistics* 7(1), pp. 217–226. doi: https://doi.org/10.7275/scil.2147. See (SCiL).

9. Damir Cavar, Ludovic V. Mompelat, Muhammad S. Abdo (2024) *The Typology of Ellipsis: A Corpus for Linguistic Analysis and Machine Learning Applications.* Pages 46–54 of Michael Hahn et al. (eds.) Proceedings of the 6$^{\text{th}}$ Workshop on Research in Computational Linguistic Typology and Multilingual NLP. Association for Computational Linguistics, St Julian's, Malta. See ACL Special Interest Group on Typology (SIGTYP) 2024, collocated with the 18th Conference of the European Chapter of the Association for Computational Linguistics.

Conference Talks 2024–2025

1. Cavar, D., Koushik Reddy Parukola, Shane Sparks (2025) *Old Wine in New Bottles: Using Classical Word Embeddings in Quantum NLP Systems.* Paper to be presented at the Midwest Speech and Language Days 2025, University of Notre Dame, April 2025.

2. A. Karkala Pai, D. Cavar (2025) *A Voice-based Detection of Parkinson's Disease: Feature Selection and Classification Results.* Paper to be presented at the Midwest Speech and Language Days 2025, University of Notre Dame, April 2025.

3. R. Shrivastava, T. Sun, S.L. Anusha Chebolu, M. Kodandapani Naidu, T. Jayaprakash, J. Decatur, A. Bajpai, R. Wang, D. Cavar (2025) *Improving LLM Reasoning Through Ontology-driven Knowledge Graphs: A Comparative Study of Generating Ontologies for Medical RAGs.* Poster to be presented at the Midwest Speech and Language Days 2025, University of Notre Dame, April 2025.

4. Damir Cavar, Koushik Reddy Parukola (2025) *Word and Text Similarity Using Classical Word Embeddings in Quantum NLP Systems.* Satellite Workshop: Quantum Machine Learning in Signal Processing and Artificial Intelligence at the 2025 IEEE International Conference on Acoustics, Speech, and Signal Processing. Hyderabad, India.

5. Entangled Meanings: Classification and Ambiguity Resolution in Near–Term QNLP. Paper presented at the Quantum AI Workshop at the IEEE Quantum Week 2024, Montreal, Canada, September 2024.

6. *Semantic Similarities using Classical Embeddings in Quantum NLP.* Poster presented at the IEEE Quantum Week 2024, Montreal, Canada, September 2024.

7. Damir Cavar, Zoran Tiganj, Ludovic Mompelat, Billy Dickson (2024) *Computing Ellipsis Constructions: Comparing Classical NLP and LLM Approaches.* Paper presented at the 2024 Meeting of the Society for Computation in Linguistics (SCiL).

8. Van Holthenrichs, Damir Cavar, Zoran Tiganj, Billy Dickson (2024) *On Ellipsis in Slavic: The Ellipsis Corpus and Natural Language Processing Results.* Paper presented at The 33rd Annual Meeting of Formal Approaches to Slavic Linguistics. Halifax, Canada.

9. *The Hoosier Ellipsis Corpus (HELC): Documenting Linguistic Dark Matter* (2024) Damir Cavar, Ludovic Mompelat, Muhammad S. Abdo. Poster presented at the Midwest Speech and Language Days at the University of Michigan in Ann Arbor, April 15-16, 2024.

10. *The Hosiers Ellipsis Corpus: Building a Corpus of Ellipsis for Arabic Natural Language Processing* (2024) Muhammad S. Abdo, Damir Cavar. Poster presented at the Midwest Speech and Language Days at the University of Michigan in Ann Arbor, April 15-16, 2024.

11. *Quantum Natural Language Processing* (QNLP) (2024) Damir Cavar, Presentation at the Quantum Day 2024 Seminar Series, organized by Quantum Technologies for Everyone (QuTE) at Indiana University Bloomington, April 14th 2024.

12. *Quantum-Natural Language Processing (NLP) and Machine Learning (ML)* (2024) Damir Cavar, Presentation at the CQT - Center for Quantum Technologies, NSF Industry/University Cooperative Research Center (IUCRC) Year 2, Phase I, Spring 2024 Industry Advisory Board Meeting, April 3-4, 2024, University of Notre Dame, South Bend, IN.

13. *Quantum Natural Language Processing and Machine Learning.* Luddy-Crane Summit on March 29, 2024 at Indiana University Bloomington.

14. *Generative AI and Knowledge Representations.* Luddy-Crane Summit on March 29, 2024 at Indiana University Bloomington.

15. *On Ellipsis in Slavic: The Ellipsis Corpus and Natural Language Processing Results.* Van Holthenrichs, Damir Cavar, Zoran Tiganj, Billy Dickson. Paper presented at the 33rd Annual Meeting of Formal Approaches to Slavic Linguistics. Halifax, Canada.

16. *The Typology of Ellipsis: A Corpus for Linguistic Analysis and Machine Learning Applications.* Damir Cavar, Ludovic V. Mompelat, Muhammad S. Abdo. Paper to be presented at the ACL Special Interest Group on Typology (SIGTYP) 2024, collocated with the 18th Conference of the European Chapter of the Association for Computational Linguistics, St Julian's, Malta.

17. *Ellipsis in Arabic: Using Machine Learning to Detect and Predict Elided Words.* Damir Cavar, Muhammad S. Abdo, and Billy Dickson. Paper presented at the Arabic Linguistic Society (ASAL) 37 Conference, February 2024, New York City.

18. *Building a Multilingual Ellipsis Corpus.* IU presentation, Luddy School, authors: Calvin Josenhans, John Phillips, Khai Willard, Luis Abrego, Yuchen Yang, Niko Kilo, Damir Cavar

Service 2024-2025

1. Member of the College Policy Committee, Spring 2024.

2. In charge of the Research Colloquium of the Department of Linguistics, Fall 2024 and Spring 2025.

3. Directing the Natural Language Processing Lab, weekly meetings during the semesters and breaks, see `https://nlp-lab.org/`.

4. Directing the Quantum NLP and AI Study Group, weekly meetings during the semesters and breaks, see `https://nlp-lab.org/quantumnlp/`.

5. Organizing the Quantum AI and NLP 2025 conference, August 2025, organizer and Conference Chair, see `https://qnlp.ai/`.