

Kara Schatz

(630)-487-6751 kmschat2@ncsu.edu go.ncsu.edu/kara-schatz linkedin.com/in/schatzkara/ github.com/schatzkara

EDUCATION

Doctorate of Philosophy in Computer Science

Expected Graduation: Fall 2024

North Carolina State University (NCSU), Raleigh, North Carolina

College of Engineering — Department of Computer Science

GPA: 4.0

Dissertation Topic: Knowledge Graph Management and Analysis for Domain Scientists

Dissertation co-advisors: Dr. Rada Chirkova and Dr. Alexander Tropsha

Master of Science in Computer Science

May 2023

North Carolina State University (NCSU), Raleigh, North Carolina

College of Engineering — Department of Computer Science

GPA: 4.0

Bachelor of Science in Mathematics and Computer Science, Minor in Statistics

May 2020

Xavier University, Cincinnati, Ohio

College of Arts and Sciences — Department of Mathematics and Department of Computer Science

University Scholar; Summa Cum Laude; GPA: 3.966

PROFESSIONAL DEVELOPMENT

Graduate Teaching Certificate

Expected: May 2024

North Carolina State University

- Working on a certificate to hone teaching skills through a series of workshops, teaching experience, and reflection.
- Workshops: Teaching College, Engaging Diverse Learners, Using Learning Outcomes, Mentoring 101, Helping Undergraduate Students with Common Research Challenges, Student Engagement Using Digital Tools, Processing Course Evaluations

Inclusive Excellence Certificate

October 2023

North Carolina State University

- Completed a certificate to increase effectiveness in creating workplaces and classrooms that exhibit inclusive excellence.
- Workshops: Applied Cultural Identity, Understanding Unconscious Bias, The Macro of Microaggressions, Fundamentals of Equal Opportunity

Graduate Writing Certificate

June 2023

North Carolina State University

- Completed a certificate to hone academic and professional writing skills through a series of workshops, writing experience, and a final work product.
- Workshops: Writer's Toolkit, How to Write Concise Sentences, How to Organize Reports, How to Craft Abstracts and Introductions, How to Synthesize Literature Reviews, How to Structure Results and Discussion Sections, How to Summarize Your Work, How to Compose Professional Emails

Reflective Educational Design Inclusive Teaching Certificate

May 2023

North Carolina State University

- Completed a certificate focused on developing resources and skills for applying inclusive course pedagogical practices that support all learners.
- Topics: Inclusive Teaching Methodology, Cultural Relevance and Responsiveness, Community and Belonging, Access and Empowerment, Awareness and Perspectives, Global Learning

TEACHING/MENTORING EXPERIENCE

Adjunct Instructor – Information Technology

October 2023 – Present

Wake Technical Community College

- Instructor of Record: Software Developer Using Python, 20-week course meeting twice a week from October to March with 24 students enrolled.
- Responsibilities include: planning and conducting lectures, creating course materials, and grading student work.
- Course is designed to prepare students for two Certiport certifications: IT Specialist Python and IT Specialist Databases.

Online Instructor

May 2023 – Present

iD Tech

- Teaching computer science and math topics to students from elementary through high school.
- Conducting online private lessons for individual students. Instructed roughly 25 different students so far on a variety of topics, including Python, Java, C++, Machine Learning and Artificial Intelligence, Data Science, Algebra, and Pre-calculus. Lessons last 1 hour for which I create my own lessons plans tailored to the student's individual needs and interests.

- Taught Online Teen Academies over the summer on Machine Learning and Artificial Intelligence using Nvidia Jetson Nanos. Each session lasted two weeks, two hours per day and had 5-10 students. Learned and taught a preexisting curriculum from Nvidia, which included an open-ended project for which I assisted students from ideation to troubleshooting to completion.

Enrichment Session Assistant

October 2022 – Present

FLAMES@MSEN Outreach Program

- Assisting with a computer science outreach enrichment program for high schoolers with one to two other volunteers.
- Teaching lessons on game design and development using Unreal Engine and GameMaker software, helping students with challenges such as ideation, technical and comprehension issues, and debugging.

Department of Computer Science Research Mentor

June 2021 – December 2021

North Carolina State University

- Mentored a rising high school senior in computer science research in the area of database and knowledge management.
- Met with the student weekly to discuss progress and next steps, prepared assignments for the student to complete, and taught necessary skills.
- Worked on a database curation project in which we performed extensive web scraping and entity-matching to curate a knowledge graph database in the biomedical domain by adding valuable metadata and partially aligning it to an existing biomedical ontology.
- Taught student about knowledge graphs, database management systems, ontologies, Python, and web scraping.

The Engineering Place Teacher Team Lead

June 2021 – August 2021

North Carolina State University

- Led a Python camp for high schoolers with two other camp leads and adapted an existing curriculum to the time constraints and goals of the camp.
- Taught the basics of Python through an application to natural language processing and sentiment analysis.
- Specifically led the Python programming lessons each day, while other camp leads let the conceptual lessons about sentimental analysis.
- Taught students with little to no programming experience and helped students individually and in groups.

Office of Academic Support Supplemental Instruction Leader

August 2018 – May 2020

Xavier University

- Supplemental Instruction (SI) leader for the Computer Science II course, which covers Java and object-oriented programming concepts, for four semesters, covering 6 class sections in total and over 100 students.
- Responsibilities included: attending class, holding 3 hour-long SI session per week, and holding 2 office hours per week.
- Assisted students with individual and group work during class, SI sessions, office hours, and via email.
- Created my own worksheets with problem sets/projects and solutions for each SI session.
- SI sessions consisted of assisting students one-on-one and in groups, reteaching concepts from class when needed, working through practice worksheets, leading exam review, helping with course assignments, and providing debugging assistance.
- Met with instructor to discuss students' trouble points, necessary changes to course materials, and other course logistics.
- Additionally held office hours for students in the Foundations of Higher Math (Discrete Math) course.

Math Tutoring Lab Tutor

August 2017 – May 2020

Xavier University

- Tutored students in various mathematics courses at Xavier with any related coursework/concepts.

Enrichment Session Leader

October 2017 – February 2020

Xavier University Xploratory

- Created lesson plans in Origami, Cryptology, and Fractals, and taught to middle schoolers at 3 sessions each year.

NOYCE Summer Program Camp Lead

June 2017 – August 2017

Xavier University

- Ran a STEM camp for elementary and middle schoolers.
- Created lessons about various STEM fields through the lens of *The Martian*/"Survival on Mars."
- Assisted with the Xavier Summer ICE Camp, a camp to spark the interest of young girls in STEM fields.

PUBLICATIONS

Pei-Yu Hou, Jing Ao, **Kara Schatz**, Alexey Gulyuk, Yaroslava Yingling, Rada Chirkova. "Provenance-Aware Data Integration and Querying for Knowledge Graphs." Manuscript accepted for publication in September 2023.

Jing Ao, **Kara Schatz**, and Rada Chirkova, "Trend Surfing: Effective and Efficient Retrieval of Unusual Temporal Trends," in *9th International Conference on Data Mining (DTMN)*, pp. 100–123, Computer Science & Information Technology (CS & IT), September 2023.

Kara Schatz, Daniel Korn, Alexander Tropsha, and Rada Chirkova, "Workflow for Domain- and Task-Sensitive Curation of Knowledge Graphs, with Use Case of DRKG," in *2022 IEEE International Conference on Big Data (Big Data)*, pp. 3692–3701, IEEE, December 2022.

Kara Schatz, Cleber Melo-Filho, Alexander Tropsha, and Rada Chirkova, “Explaining Drug-Discovery Hypotheses Using Knowledge-Graph Patterns,” in *2021 IEEE International Conference on Big Data (Big Data)*, pp. 3709–3716, IEEE, December 2021.

Kara Schatz, Erik Quintanilla, Shruti Vyas, and Yogesh Rawat, “A Recurrent Transformer Network for Novel View Action Synthesis,” in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVII 16*, pp. 410–426, Springer International Publishing, August 2020.

MANUSCRIPTS

Kara Schatz, Pei-Yu Hou, Alexey Gulyuk, Yaroslava Yingling, Rada Chirkova. “BUILD-KG: Integrating Heterogeneous Data Into Analytics-Enabling Knowledge Graphs.” Manuscript submitted for publication in October 2023.

Kara Schatz, Jon-Michael Beasley, Cleber Melo-Filho, Alexander Tropsha, and Rada Chirkova. “Using Extraction and Evaluation of Explanations for Drug-Repurposing on Knowledge Graphs.” Manuscript submitted for publication in October 2023.

Kara Schatz, Alexander Tropsha, Rada Chirkova. “GAME: Improving Efficiency and Effectiveness of Knowledge-Graph Rule Mining via Data Reduction.” Manuscript submitted for publication in September 2023.

Daniel Korn, Pei-Yu Hou, **Kara Schatz**, Jon-Michael Beasley, Rada Chirkova, and Alexander Tropsha. “Towards Improving the Efficiency of Drug Repurposing by Leveraging Node Promiscuity in Biomedical Knowledge Graphs.” Manuscript submitted for publication in May 2023.

PRESENTATIONS

9th International Conference on Data Mining (DTMN)

September 16 – 17, 2023

Virtual

- Virtual conference presentation for DTMN 2023 consisting of a live presentation and a Q&A session.
- J. Ao, **Kara Schatz**, and R. Chirkova, “Trend Surfing: Effective and Efficient Retrieval of Unusual Temporal Trends,” in *DTMN*, pp. 100–123, CS & IT, 2023.

NSF STEPS Science and Technologies Center Annual Site Visit Poster Session

May 31, 2023

North Carolina State University

- Poster session for NSF evaluators.
- Poster covered work done on the *Convergence Informatics Project* for STEPS.

IEEE Big Data Conference

December 17 – 20, 2022

Virtual

- Virtual conference presentation for IEEE Big Data 2022 consisting of a video, a live presentation, and a Q&A session.
- **Kara Schatz**, D. Korn, A. Tropsha, and R. Chirkova, “Workflow for Domain- and Task-Sensitive Curation of Knowledge Graphs, with Use Case of DRKG,” in *2022 IEEE Big Data*, pp. 3692–3701, IEEE, 2022.

NCSU Graduate Student Research Symposium

April 6, 2022

North Carolina State University

- Nominated by my department to present a poster at the 15th Annual Symposium to showcase NCSU’s graduate-level research.
- Poster covered work done on the *Explaining Hypotheses on Knowledge Graphs Project*.

IEEE Big Data Conference

December 15 – 18, 2021

Virtual

- Virtual conference presentation for IEEE Big Data 2021 consisting of a video, a live presentation, and a Q&A session.
- **Kara Schatz**, C. Melo-Filho, A. Tropsha, and R. Chirkova, “Explaining Drug-Discovery Hypotheses Using Knowledge-Graph Patterns,” in *2021 IEEE Big Data*, pp. 3709–3716, IEEE, 2021.

European Conference on Computer Vision

August 23 – 28, 2020

Virtual

- Virtual poster session for ECCV 2020 consisting of videos and Q&A sessions.
- **Kara Schatz**, E. Quintanilla, S. Vyas, and Y. S. Rawat, “A Recurrent Transformer Network for Novel View Action Synthesis,” in *ECCV 2020*, pp. 410–426, Springer International Publishing, 2020.

Pi Mu Epsilon Student Talks

September 21, 2018

Miami University

- Presented 2018 REU research on estimating bacterial recombination rates to students and faculty in the Ohio Theta Chapter of Pi Mu Epsilon as well as other guests.
- Poster covered work done on the *Modeling Bacterial Recombination Rates Project*.

MathFest Student Paper Session

August 1 – 4, 2018

Denver, Colorado

- Presented 2018 REU research on estimating bacterial recombination rates to other undergraduate students and mathematicians at MathFest.
- Poster covered work done on the *Modeling Bacterial Recombination Rates Project*.

RESEARCH EXPERIENCE

NCATS Biomedical Data Translator: ARAGORN

November 2020 – Present

Renaissance Computing Institute (RENCI)

- **Computational Drug-Repurposing Project:** Collaborating with researchers on the ARAGORN team at RENCI, who developed the large-scale biomedical graph called ROBOKOP, to computationally search the graph for potentially promising drug-repurposing hypotheses, or drug-disease relationships that may lead to treatment development after confirming clinical trials. Working specifically on knowledge graph inference rule generation for ARAGORN's "creative-mode" queries, which enable users to query a node pair of interest to them; if no relationship between them exists in ROBOKOP, then inference rules are applied in succession to identify any possible inferred connection between the node pair. The ARAGORN team has begun collaborations with the non-profit Every Cure, which will bring the drug-repurposing hypotheses generated to clinical trials for further drug development.

Department of Computer Science

August 2020 – Present

North Carolina State University

- **Knowledge Graph Reduction Project:** Developed an approach to abstract, or compress, knowledge graphs in such a way that very little information is lost. In this way, the abstract knowledge graph can be used for more efficient analytics, while not sacrificing much of the effectiveness of the analytics or their results. Tested knowledge graph inference rule mining on the compressed graphs and found that the mining stage is significantly faster, but very few inference rules are lost on the compressed graphs. Submitted a manuscript of this work for publication.
- **Iterative Metapath Mining Project:** Developed an approach to mine metapaths from knowledge graphs in a ping-pong style with domain scientists as humans-in-the-loop. The approach uses pairs of related and unrelated nodes and mine metapaths that bring the related nodes together in latent embedding space, while driving the unrelated nodes apart. In this way, the metapaths discovered provide some semantic reasoning about the relation between the node pairs. Manuscript of this work in preparation.
- **Provenance-Aware Data Integration Project:** Developed a domain-agnostic approach to integrate data across multiple granularity levels: standard-level data and summary-level data. During the integration the appropriate standard-level data that contributes to the summary-level data is identified and maintained as provenance to enable efficient and effective querying of both the data and provenance after integration. Specifically worked on the problem formalism and well as helping to develop the approach. A paper written about this work has been accepted for publication.
- **Node Promiscuity Project:** Developed the notion of promiscuity in knowledge graphs, which is related to node degree and connectivity. Devised an algorithm for finding least promiscuous paths in knowledge graphs, which have been found to be more interesting to domain scientists than promiscuous paths. Specifically worked on helping to create the algorithm and worked heavily on the formalism of promiscuity and promiscuous path-searching. Submitted a manuscript of this work for publication.
- **Knowledge Graph Curation Project:** Refined and improved the "Drug Repurposing Knowledge Graph" released by AWS. Previously this knowledge graph consisted only of complex identifiers connected by relationships. Through web-scraping and database cross-referencing, we have refined all the semantic triples to now include clear entity names, relationship types, and other desired attributes. Overall, this makes the knowledge graph more usable for analysis, especially for biomedical researchers and users, since the data can be understood. Published a workflow describing our curation steps in the *2022 IEEE International Conference on Big Data*, and made a presentation about this project at the conference.
- **Explaining Hypotheses on Knowledge Graphs Project:** Designed a domain-agnostic approach for extracting and evaluating explanations from knowledge graphs for proposed hypotheses. The approach derives explanations via backward-chaining using knowledge graph triples and inference rules to justify the truth of the hypotheses. The explanation-evaluation metrics offer a global evaluation of each explanation based on existing evidence in the knowledge graph. Tested the approach on the biomedical application of explaining and evaluating drug-repurposing hypotheses, and collaborated with biomedical experts to study the domain reasonableness of our output. Published this work in the *2021 IEEE International Conference on Big Data (Big Data)*, and made a presentation about this project at the conference. Also presented a poster about this work at the 15th Annual NCSU Graduate Student Research Symposium. Submitted an extended manuscript of this work for publication.
- **Identifying Unusual Temporal Trends Project:** Developed a heuristic method to efficiently and effectively identify outlier temporal trends in data sets. Specifically worked on the theoretical justification for our approach, which involved multiple time-complexity proofs, as well as deriving a formula to estimate user efforts in unassisted searches of the search space for temporal trends. Published this work in the *9th International Conference on Data Mining (DTMN)*, and made a presentation about this project at the conference. An extended manuscript of this work is in preparation.

Science and Technologies for Phosphorus Sustainability (STEPS) Center

August 2021 – August 2023

North Carolina State University

- **Convergence Informatics Project:** Designed a system for uniformly compiling all scales and types of STEPS data into a knowledge graph. Led the efforts to build the knowledge graph from scratch, and design its ontology and schema as well. Used Python and Neo4j to build the knowledge graph. Through close collaboration with STEPS domain experts, designed an appropriate ontology and ingested several data types accordingly. Presented a poster on this work during the NSF STEPS Annual Site Visit. Submitted a manuscript of this work for publication.

Department of Computer Science

January 2020 – May 2020

Xavier University

- **Computer Science Thesis Project:** Developed an AI to play the popular phone and web game 2048, under the advisement of Dr. Elizabeth Johnson. Developed a state space for the game and implemented AI search techniques with several different heuristics to determine the best strategy. Implemented the project, including a GUI, in Python. Presented this project during the department's Senior Capstone Event.

Department of Mathematics

August 2019 – May 2020

Xavier University

- **Mathematics Senior Thesis Project:** Derived and proved original formulas for the linear bandwidth and linear bandwidth criticality properties of complete bipartite and complete tripartite graphs, under the advisement of Dr. Eric Bucher. Determining the linear bandwidth of graphs is an NP-complete problem, except for certain classes of graphs, such as those explored in this project. Wrote a technical report about this work and presented it during the department's Senior Capstone Event.

Center for Research in Computer Vision

May 2019 – February 2020

University of Central Florida

- Part of a Research Experiences for Undergraduates (REU) Program.
- **Novel View Action Synthesis Project:** Developed a machine learning framework to solve the novel view action synthesis problem, which had little prior research at the time, under the advisement of Dr. Yogesh Rawat and Dr. Mubarak Shah. The model was a recurrent transformer network that converting an action video to a new viewpoint by using a single image prior from the new view. Implemented this framework in Python using the Pytorch library and conducted experiments on the large-scale multi-view action recognition NTU-RGB+D dataset. Continued this research for several months after the REU ended. Published this work in the *16th European Conference on Computer Vision (ECCV)*, and made a presentation about this project at the conference.

Hello Research Workshop

October 26 – 28, 2018

Indiana University

- Part of an Opportunities for Undergraduate Research in Computer Science (OurCS) Program.
- **Human Identification Project:** Explored machine learning and its applications by developing an approach to identify humans based on audio and video data. Created a proof-of-concept model trained on audio and video feed of members within our group. Presented this work as a group at the end of the workshop.

Department of Mathematics and Statistics and Department of Biology

May 2018 – July 2018

University of North Carolina Greensboro

- Part of a Research Experiences for Undergraduates (REU) Program.
- **Modeling Bacterial Recombination Rates Project:** Created mathematical models and computer simulations to estimate recombination rates in bacteria, under the advisement of Dr. Louis-Marie Bobay and Dr. Jonathan Rowell. This project involved deriving formulas for bacteria recombination rates using probability and combinatorics. Also programmed a simulation in Python to model bacterial recombination based on genetics principles. Presented this research at 2 events: the MathFest Student Paper Session and the Pi Mu Epsilon Student Talks.

OTHER WORK EXPERIENCE**Mathematics Department Assistant**

August 2017 – May 2020

Xavier University

- Graded homework assignments for various courses, e.g., Discrete Math, Languages and Automata, and Calculus II.

SKILLS AND RELEVANT COURSEWORK**Programming Languages:** Python, Java, C++, C, SQL, Cypher, MIPS, Scheme, MATLAB, Golang**Programming Tools:** Git, Neo4j, Pytorch, Google Cloud**Graduate Courses:** Computational Logic, Design and Analysis of Algorithms, Databases, Artificial Intelligence, Graph Data Mining, Graph Theory, Technical Communication for Engineering Research, Algorithms on Strings, Human-Computer Interaction, Automated Learning and Data Analysis**Undergraduate Courses:** Data Structures and Algorithms, Machine Organization, Programming Languages, Databases, Artificial Intelligence, Operating Systems, Networking, Security, Graph Theory, Data Mining**AWARDS****NCSU College of Engineering Graduate Enhancement Award**

April 2023

- Awarded to a small number of outstanding graduate students.
- Award amount: \$2,850

NCSU Department of Computer Science Outstanding Research Award Nomination

February 2023

- Recognizes a doctoral student for outstanding research contributions in the previous calendar year.

NCSU Provost's Doctoral Fellowship

August 2020 – July 2021

- Awarded to support outstanding doctoral students entering the program.
- Award amount: \$30,000

Clare Boothe Luce Scholarship Recipient

May 2018 – May 2020

- Awarded to an undergraduate woman in a STEM field who embodies the characteristics of the Clare Boothe Luce legacy and plans to attend graduate school in her STEM field.

John F. Niehaus Award

April 2020

- Given in memory of John F. Niehaus to an outstanding senior majoring in computer science.

Kramer-Miller Mathematics Award

April 2019 & April 2020

- Presented to a senior Mathematics Major for having achieved great distinction in the study of mathematics.

Consortium for Mathematics and its Applications Meritorious Award

February 2018

- Received this award for my team's submission in the Consortium for Mathematics and its Applications Interdisciplinary Contest in Modeling.
- Scored in the top 11% of all submissions for this problem.

SERVICE ACTIVITIES

3rd Workshop on Knowledge Graphs and Big Data in conjunction with IEEE Big Data 2023 <i>Program Committee Member</i>	December 2023
FLAMES@MSEN Outreach Program <i>Volunteer Enrichment Session Assistant</i>	October 2022 – Present
ACM Transactions on Knowledge Discovery from Data <i>Submission Reviewer</i>	May 2023
DiamondHacks Hack-a-thon <i>Volunteer Mentor</i>	March 26 – 27, 2022
46th International Conference on Very Large Databases (VLDB) <i>Student Volunteer</i>	August 31 – September 4, 2020
Xavier Xploratory <i>Volunteer Enrichment Session Leader</i>	October 2017 – February 2020

OTHER ACTIVITIES

NCSU STARS Computing Corps <i>Member</i>	August 2021 – Present
NCSU Women in Computer Science <i>Member</i>	August 2020 – Present
Xavier University Math Club <i>President</i>	August 2018 – May 2020
<ul style="list-style-type: none"> Organized events where members can bond with other students who share similar interests. Before August 2018, the Math Club at Xavier had not held a meeting in over 2 years. By 2020, we had more than 20 active members attending our events. 	
Xavier University Women in Computing <i>Member</i>	August 2018 – May 2020
Xavier University Computer Science Club <i>Member</i>	August 2018 – May 2020

PROFESSIONAL AFFILIATIONS

Pi Mu Epsilon (Mathematical Honor Society) <i>Member of the Ohio Theta Chapter</i>	Inducted November 2017
Phi Beta Kappa (Academic Honor Society) <i>Member of the Pi of Ohio Chapter</i>	Inducted April 2020

COMPETITIONS

William Lowell Putnam Mathematical Competition <i>Participant</i>	December 1, 2018
Association for Computing Machinery – International Collegiate Programming Contest <i>Team Member</i>	November 10, 2018
Consortium for Mathematics and its Applications Interdisciplinary Contest in Modeling <i>Team Member</i>	February 8 – 12, 2018