

HIMARSHA R. JAYANETTI

Ph.D. Student, Computer Science, Old Dominion University, Norfolk, VA

@hjaya002@odu.edu | (757) 822 0108 | in himarshaj | himarshaj | himarshaj.github.io | @HimarshaJ

SUMMARY

- **PhD Candidate:** Computer Science, Expected graduation date: May 2027.
- **Publications:** 8 conference papers, 2 journal articles, 5 student conference/poster papers (Google Scholar).
- **Awards:** Best Student Paper (TPDL '22), Best Short Paper (JCDL '23), Best Short Paper Nominee (JCDL '23), and Best Paper Data Science Track and Best Paper Overall (MSVSCC '23 Student Conference)
- **Travel Grants:** Received over \$4,000 in competitive travel awards and conference support.
- **Academic Service:** Served on 5 program committees and reviewer roles for conferences including JCDL, TPDL, and CIKM (2021–2026).
- **Internship:** Department of National Archives, Sri Lanka (2016).
- **Industry Experience:** Network Support Engineer at Exetel Telecommunications, Sri Lanka (2017–2019).
- **Internship:** Los Alamos National Laboratory (2022).
- **Internship:** Internet Archive, as a Google Summer of Code Contributor (2025).

RESEARCH INTERESTS

Web Archiving, Digital Libraries, Social Media, Web Science, Data Science, Machine Learning, Information Retrieval, Computational Social Science, Human-Centered AI

EDUCATION

Ph.D., Computer Science, Old Dominion University

📅 Fall 2020 – Present

📍 Norfolk, Virginia, USA

M.S., Computer Science, Old Dominion University

📅 Fall 2019 – Summer 2023

📍 Norfolk, Virginia, USA

B.Eng., Computer Science, Gujarat Technological University

📅 August 2013 – May 2017

📍 Ahmadabad, Gujarat, India

WORK EXPERIENCE

Graduate Research / Teaching Assistant, Old Dominion University

📅 Aug 2019 – Present

📍 Norfolk, Virginia, USA

- **Research Assistant:** Conduct research in web archiving, social media analysis, data science, and information retrieval (Spring 2021 – Present, excluding Spring 2026).
- **Teaching Assistant:** CS 120 (Fall 2019; Spring & Fall 2020), **Instructor:** CS 450/550 (Spring 2026).

Google Summer of Code Contributor, Internet Archive

📅 June 2025 – Aug 2025

📍 Remote

- Developed an LLM-based model for detecting social media content in TV news via logo and post screenshot detection.
- This was a summer internship at the Internet Archive as a Google Summer of Code contributor.

Intern, Los Alamos National Laboratory

📅 June 2022 – Aug 2022

📍 Remote

- Developed a system to generate social cards for visualizing scientific PDFs.
-

Network Support Engineer, Exetel Telecommunications

📅 July 2017 – July 2019

📍 Colombo, Sri Lanka

- Worked with ISP infrastructure and troubleshooting Australia's broadband technologies.
-

Intern, Department of National Archives

📅 Jan 2016 – Dec 2016

📍 Colombo, Sri Lanka

- Worked on desktop applications, web design, networking, and hardware.
-

PUBLICATIONS

- Jayanetti, Himarsha R, Kritika Garg, Sawood Alam, Michael L Nelson, and Michele C Weigle (2024). "Robots still outnumber humans in web archives in 2019, but less than in 2015 and 2012". In: *International Journal on Digital Libraries*, pp. 1–17.
- Jayanetti, Himarsha R (2024). "Evaluating Social Media Reach via Mainstream Media Discourse". In: *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, pp. 5455–5458.
- Jones, Shawn M, Himarsha R Jayanetti, Martin Klein, Michele C Weigle, and Michael L Nelson (2023). "Synthesizing Web Archive Collections into Big Data: Lessons from Mining Data from Web Archives". In: *International Conference on Theory and Practice of Digital Libraries*. Springer, pp. 220–229.
- Jayanetti, Himarsha R, Erika Frydenlund, and Michele C Weigle (2023a). "Exploring Xenophobic Events through GDELT Data Analysis". In: *arXiv*. DOI: [10.48550/arXiv.2305.01708](https://doi.org/10.48550/arXiv.2305.01708). Modeling, Simulation and Visualization Student Capstone Conference (MSVSCC '23), ODU-VMASC.
- Jayanetti, Himarsha R, Erika Frydenlund, and Michele C Weigle (2023b). "Xenophobic Events vs. Refugee Population—Using GDELT to Identify Countries with Disproportionate Coverage". In: *arXiv*. DOI: [10.48550/arXiv.2308.05038](https://doi.org/10.48550/arXiv.2308.05038). 16th International Conference on Social Computing, Behavioral-Cultural Modeling & Prediction and Behavior Representation in Modeling and Simulation (SBP-BRIMS 2023).
- Garg, Kritika, Himarsha R Jayanetti, Sawood Alam, Michele C Weigle, and Michael L Nelson (2023). "Challenges in replaying archived Twitter pages". In: *International Journal on Digital Libraries* 25.2, pp. 217–236.
- Choudhury, Muntabir Hasan, Lamia Salsabil, Himarsha R Jayanetti, Jian Wu, William A Ingram, and Edward A Fox (2023). "MetaEnhance: Metadata Quality Improvement for Electronic Theses and Dissertations of University Libraries". In: *2023 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. IEEE, pp. 61–65.
- Bragg, Haley, Himarsha R Jayanetti, Michael L Nelson, and Michele C Weigle (2023). "Less than 4% of archived Instagram account pages for the disinformation dozen are replayable". In: *2023 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. IEEE, pp. 102–106.
- Jones, Shawn Morgan, Himarsha R Jayanetti, Alex Osborne, Paul Koerbin, Klein Martin, Michele C Weigle, and Michael L Nelson (2022). "The DSA toolkit shines light into Dark and Stormy Archives". In: *Code {4} Lib* 53.
- Jayanetti, Himarsha R, Shawn M Jones, Martin Klein, Alex Osbourne, Paul Koerbin, Michael L Nelson, and Michele C Weigle (2022). "Creating structure in web archives with collections: different concepts from web archivists". In: *International Conference on Theory and Practice of Digital Libraries*. Springer, pp. 450–458.
- Jayanetti, Himarsha R, Kritika Garg, Sawood Alam, Michael L Nelson, and Michele C Weigle (2022). "Robots still outnumber humans in web archives, but less than before". In: *International Conference on Theory and Practice of Digital Libraries*. Springer, pp. 245–259.
- Garg, Kritika, Himarsha R Jayanetti, Sawood Alam, Michele C Weigle, and Michael L Nelson (2022). "Caching HTTP 404 Responses Eliminates Unnecessary Archival Replay Requests". In: *International Conference on Asian Digital Libraries*. Springer, pp. 329–344.

- Garg, Kritika, Himarsha R Jayanetti, Sawood Alam, Michele C Weigle, and Michael L Nelson (2021). "Replaying Archived Twitter: When your bird is broken, will it bring you down?" In: *2021 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. IEEE, pp. 160–169.
- Choudhury, Muntabir Hasan, Himarsha R Jayanetti, Jian Wu, William A Ingram, and Edward A Fox (2021). "Automatic metadata extraction incorporating visual features from scanned electronic theses and dissertations". In: *2021 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. IEEE, pp. 230–233.






In Review

- Jayanetti, Himarsha R., Sivakanesan Dhanushkanda, Shuai Hao, Michael L. Nelson, and Michele C. Weigle (2026). "Nobody Truly Agrees on Sentiment: Humans, Bespoke Tools, and LLMs Struggle with Social Media Texts". In: *A Peer-Reviewed Venue*. [Short Paper] (In Review).
- Jayanetti, Himarsha R., Sawood Alam, Alexander C. Nwala, Michael L. Nelson, and Michele C. Weigle (2026). "A Multimodal LLM-Based Framework for Detecting Social Media Content in TV News Archives". In: *A Peer-Reviewed Venue*. [Full Paper] (In Review).

ACHIEVEMENTS

-  **ODU Computer Science Graduate Society (CSGS) Hackathon Winner**
(2025) Won first place as part of a three-member team ([blog](#)).
-  **Best Short Paper Award**
(2023) Short Paper at ACM/IEEE Joint Conference on Digital Libraries.
-  **Best Overall Paper and Best Paper in Data Science track Award**
(2023) Full Paper at the 16th Annual Modeling, Simulation, and Visualization Student Capstone Conference, Prize: \$500 (\$350 Best Overall Paper + \$150 Best Data Science Track Paper).
-  **Best Student Paper**
(2022) Full Paper at Theory and Practice of Digital Libraries Conference, Prize: €500.
-  **Irwin B. Levinstein Scholarship**
(2021) Old Dominion University, Funding: \$2500.
-  **First Class Honors**
(2017) B.Eng., Computer Science, Gujarat Technological University, India.
-  **Nehru Memorial Scholarship**
(2013) Indian Council for Cultural Relations and University Grants Commission of Sri Lanka.

TRAVEL GRANTS

-  **CAPWIC Scholarship**
(2025) Funding covering conference registration (\$50) for Washington, DC.
-  **NSF Travel Grant for CIKM**
(2024) Funding: \$1250 for travel to Boise, Idaho.
-  **SBP-BRIMS Conference Scholarship**
(2023) Funding: \$1000 for travel to Pittsburgh, Pennsylvania.
-  **ACM SIGIR Student Travel Grant for JCDL**
(2023) Funding covering conference registration and \$1420 travel support for Santa Fe, New Mexico.
-  **ODU SEES Travel Award for TPD**
(2022) Funding: \$500 for travel to Padua, Italy.

GRANT PROPOSALS

- I contributed to a grant proposal extending my dissertation research on how social media content extends beyond its native platform and is integrated into television news. This proposal is currently under review.

TEACHING EXPERIENCE

- **Instructor of Record**, CS 450/550: Database Concepts (Spring 2026)
- **Teaching Assistant**, CS 120: Introduction to Information Literacy and Research (Fall 2019; Spring & Fall 2020)

SKILLS

Languages: Python, C/C++, C#, Java, JavaScript, R, Bash / shell scripting

Frameworks & Tools: Django, NodeJS, React, Docker, Elasticsearch, Kibana, GitHub

Data Science: TensorFlow, WEKA, NumPy, Pandas, NetworkX

Visualization: D3, Vega-lite, Tableau, ObservableHQ

Databases: MongoDB, MySQL, PostgreSQL, SQLite

Systems: Linux / Unix, Cloud platforms (AWS, GCP, Azure), Networking fundamentals (TCP/IP, HTTP, DNS)

Version Control: Git

Machine Learning & AI: NLP, Computer vision, Model evaluation, LLM prompting, Multimodal models

Soft Skills: Technical writing, Research communication, Cross-disciplinary collaboration, Grant writing, Project leadership, Peer code review & mentorship

SERVICES & OUTREACH

- Program Committee Member, CIKM 2026 (Full Research, Short Paper)
- Program Committee Member, JCDL 2026 (Full or Short Research Papers)
- Program Committee Member, TPDJ 2026 (Full Papers, Short and Late Breaking Results Papers)
- Program Committee Member, JCDL 2025 (Research Paper)
- Secretary, Sri Lankan Student Association at ODU (2023-2024)
- Student Mentor, 2022 NSF Research Experiences for Undergraduates (REU) program.
- President, ACM-W at ODU (2022)
- Secretary, ACM-W at ODU (2021)
- Peer Reviewer, TPDJ (2021)
- Instructor, Data Science Summer Camp (2021)
- Volunteer, STRS Summer Camp (2020)

PROJECTS

Ph.D. Dissertation

📅 2023 – Present

This study examines how social media content extends beyond its platform and is integrated into television news. Existing research largely measures impact through intra-platform engagement, overlooking television's role as a widely consumed and more trusted medium with a distinct audience. As a result, additional reach and influence of social media content beyond its native platform remain significantly underestimated. To address this gap, this work proposes a framework for detecting and analyzing social media references in TV news broadcasts.

What's Missing? Innovating Interdisciplinary Methods for Hard-to-Reach Environments

📅 2022 – 2025

Large interdisciplinary research project which was funded by the U.S. Department of Defense Minerva Research Initiative investigating methods for studying hard-to-reach environments (Khayelitsha Site-C, South Africa and Villa Caracas, Colombia). The project brought together an international team across the U.S., Canada, Norway, and Colombia. The work examines methodological and epistemological gaps when only certain research approaches are feasible. Our component focuses on analyzing these sites using public data sources such as global news and social media, alongside complementary methods including surveys, visual sociology, and citizen science across collaborating teams. The project was later discontinued before completion.

Data Science for Social Good Project

📅 2022 – 2023

This project is titled "Data Science for Social Good: Mining and Visualizing Worldwide News to Monitor Xenophobic Violence". Here, I explored xenophobic events related to refugees and migration using the GDELT 2.0 news database and APIs through visualizations. This project was funded by the 2022-2023 ODU Data Science Seed Funding Program.

Analyzing Unnecessary Traffic In Web Archives

📅 2022

The network activity on different archived web pages are examined while investigating the cause for these kinds of unnecessary requests.

Access Patterns of Robots and Humans in Web Archives

📅 2022

To identify user sessions in the access logs, the Internet Archive (IA) Wayback Machine and Portuguese Web Archive access logs are used. Based on their browsing behavior, the sessions are then categorized as human or robot.

ETD Data Enhancement

📅 2022

Improving quality of large ETD corpus using OCR and extraction tools.

Adverse Effects of Twitter's UI change on Web Archives

📅 2021

Documenting and analyzing the problems in archiving Twitter ever since Twitter forced the use of its new UI in June 2020. This study was done using the personal Twitter account of the 45th President of the United States, @realDonaldTrump, which was suspended by Twitter on January 8, 2021.

DSA Project

📅 2021

This project provides storytelling solutions to improve the understanding of web archive collections. I have contributed to several software components of the DSA project including the AIU, Raintale, and MementoEmbed. <https://oduwsdl.github.io/dsa/>.

Extracting Metadata from Scanned ETDs

📅 2020

A conditional random field (CRF) model that combines text-based and visual features is used to extract metadata from scanned ETDs. The results were compared against the previous heuristic-based method.

M.S. Thesis

📅 2019 - 2023

My master's thesis is titled "Supporting Account-based Queries for Archived Instagram Posts". There is a lot of Instagram web content preserved in web archives that the user is unable to discover since the URL of the resource (post URL) is not known. The thesis proposed several approaches to simplify the process for users interested in locating all archived Instagram posts associated with a specific user account. On April 13, 2023, I successfully defended my MS thesis.

ARTICLES

- **Summer Project as a Google Summer of Code (GSoC) Contributor**
<https://ws-dl.blogspot.com/2025/09/2025-09-29-summer-project-as-google.html>
- **A Battle of Opinions: Tools vs. Humans (and Humans vs. Humans) in Sentiment Analysis**
https://ws-dl.blogspot.com/2025/03/2025-03-26-battle-of-opinions-tools-vs_61.html
- **Tools for Using the Internet Archive's TV News Archive as Research Data**
<https://ws-dl.blogspot.com/2024/08/2024-08-07-tools-for-using-internet.html>
- **PDFServer - Our Summer Internship at LANL**
<https://ws-dl.blogspot.com/2022/12/2022-12-21-pdfserver-our-summer.html>
- **Creation Time and Published Time Are Not the Same: Estimating the Instagram Epoch**
<https://ws-dl.blogspot.com/2021/02/2021-02-20-creation-time-and-published.html>
- **Twitter Rewrites Your URLs, but Assumes You'll Never Rewrite Theirs: More Problems Replaying Archived Twitter**
<https://ws-dl.blogspot.com/2021/01/2020-01-22-twitter-rewrites-your-urls.html>
- **Twitter Added Labels On Its Old User Interface**
<https://ws-dl.blogspot.com/2020/12/2020-12-08-twitter-added-labels-on-its.html>
- **How Well Is Instagram Archived?**
<https://ws-dl.blogspot.com/2020/11/2020-11-04-how-well-is-instagram.html>
- **New Twitter UI: Replaying Archived Twitter Pages That Never Existed**
<https://ws-dl.blogspot.com/2020/11/2020-11-04-new-twitter-ui-replaying.html>
- **Twitter Was Already Difficult To Archive, Now It's Worse!**
<https://ws-dl.blogspot.com/2020/07/2020-07-15-twitter-was-already.html>