HIMARSHA R. JAYANETTI

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

📞 (757) 822 0108   @ hjaya002@odu.edu   🌐 himarshaj.github.io
🐦 @HimarshaJ   in himarshaj   ○ himarshaj

WORK EXPERIENCE

Graduate Research / Teaching Assistant @ Old Dominion University
📅 Aug 2019–Present   📍 Norfolk, VA

Teaching Assistant - I was a teaching assistant for CS 120: Introduction to Information Literacy and Research (Fall 2019, Spring 2020, and Fall 2020).
Research Assistant - I am working as a research assistant with ODU and ODURF while pursuing my M.S. and Ph.D. degrees at ODU, where I worked on different Web Archiving, Social Media, Machine Learning, and Web-related research and development projects (Fall 2019 - Present).

Network Support Engineer @ Exetel Telecommunications (Pvt) Ltd
📅 July 2017–July, 2019   📍 Colombo, Sri Lanka

I gained a thorough understanding of an ISP's telecommunications systems and infrastructure. I directly communicated with Australia's largest wholesale suppliers, including Telstra, Optus, AAPT, Opticomm, and NBN Co., about onsite related diagnosing. I could develop skills in troubleshooting ADSL, ADSL2+, bundled services, VDSL, Fiber Optic services (FTTP, HFC), VOIP, and mobile broadband difficulties for residential and small-to-medium business customers.

Department of National Archives - Intern @ Department of National Archives
📅 Jan 2016–Dec 2018   📍 Colombo, Sri Lanka

My internship training included Desktop application development, Web design, Networking and Hardware engineering.

RECENT PROJECTS

ETD Data Enhancement and Quality Improvement @ Mining ETDs
📅 Spring 2022

To enhance the data and improve the quality of a 500,000 ETD corpus. Exploring tools that allow figure and table extraction along with the captions from the ETDs and explore more satisfactory open-source OCR tools.

Analyzing Unnecessary Traffic In Web Archives
📅 Spring 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
📅 Spring 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.

DSA @ Dark and Stormy Archives
📅 Jan 2021–Dec 2021   📍 https://oduwsdl.github.io/dsa/

Developed to provide 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.

Adverse Effects of Twitter's UI change on Web Archives @ Replaying Archived Twitter
📅 Fall 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.

Extracting Metadata from Scanned ETDs @ Mining ETDs
📅 Summer 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
📅 Fall 2019–Present

There are lots of things related to Instagram in the Internet Archive's Wayback Machine that we can't discover since we are unaware of the URLs. This project aims to make it easier for interested users to find archived Instagram pages.

ABOUT ME

I'm a PhD student and a self-motivated researcher looking for an internship opportunity. Seeking to leverage my technical and communications skills in the computational sciences graduate internship program.

INTERESTS

Web Archiving · Web Science · Social Media · Digital Preservation · Data Science · Machine Learning · Information Retrieval

EDUCATION

Ph.D. (Computer Science) @ Old Dominion University, Norfolk, VA
📅 2020–Present

M.S. (Computer Science) @ Old Dominion University, Norfolk, VA
📅 2019–Present

B.Eng (Computer Science) @ Gujarat Technological University, India
📅 2013–2017

ACHIEVEMENTS AWARDS

🏆 **Irwin B. Levinstein Scholarship** (2021) Old Dominion University, Funding: $2500

🏆 **Graduated with First Class Honors** (2017) Gujarat Technological University, India

🏆 **Nehru Memorial Scholarship** (2013) Indian Council for Cultural Relations and University Grants Commission of Sri Lanka

SKILLS

Python · C/C++ · C# · Java · JavaScript · R · MongoDB · MySQL · PostgreSQL · Wooey · Django · NodeJS · React · Docker · Elasticsearch · Kibana · TensorFlow · WEKA · NumPy · Pandas · NetworkX · Seaborn · Vega-lite · D3 · Tableau · ObservableHQ · Unix · AWS

OUTREACH

🖐 **President, ACM-W** *President of the Association for Computing Machinery for Women at ODU CS (Spring, 2022))*

🖐 **Secretary, ACM-W** *Secretary of the Association for Computing Machinery for Women at ODU CS (Fall, 2021))*

🖐 **Participation, Trick or Research** *Participation at the Trick or Research event at ODU CS (2019,2020,2021))*

🖐 **Peer Reviewer** *TPDL (2021)*

🖐 **Instructor, Data Science Summer Camp** *An intensive training program to prepare high school students to work with data (2021)*

🖐 **Volunteer, STRS Summer Camp** *Teaching Assistant for Student ThinSat Research Summer Camp for high school students (2020)*

PUBLICATIONS

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.

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.

ARTICLES

- **Twitter Was Already Difficult To Archive, Now It's Worse!**
  https://ws-dl.blogspot.com/2020/07/2020-07-15-twitter-was-already.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 Added Labels On Its Old User Interface**
  https://ws-dl.blogspot.com/2020/12/2020-12-08-twitter-added-labels-on-its.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

- **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

- **Data Science High School Summer Camp 2021**
  https://ws-dl.blogspot.com/2021/09/2021-09-xx-data-science-high-school.html

- **Digging Up a Gem Through the Web Archives**
  https://ws-dl.blogspot.com/2021/09/2021-09-20-digging-up-gem-through-web.html

PROFESSIONAL MEMBERSHIPS

- **Acquired IESL Membership** 2019, Institute of Engineers Sri Lanka
  https://iesl.lk/

- **ACM/IEEE Member** 2019-Present, National student member

OTHER ACCOMPLISHMENTS

Featured in ODU CS Video Interview @ YouTube

📅 Jan, 2021

I got an opportunity to be interviewed by the College of Sciences on my research and experiences at the ODU Computer Science Department. This video was also featured on the college of science home page.

Internet Archive Highlighting #InternetArchive25 Article @ Twitter

📅 Sep, 2021

My article, "Digging Up a Gem Through the Web Archives" got highlighted and retweeted by the Internet Archive. This blog post has been viewed more than 950 times.

Archiving Twitter Project Highlighted on DSHR Blog @ First DSHR Mention, Second DSHR Mention

📅 Sep, 2021

The research project on archiving Twitter got highlighted twice on David S. H. Rosenthal's blog (DSHR Blog), a popular computer scientist within the computing community.

Archives-Unleashed Datathon - Montreal, QB, Canada @ Archives Unleashed Project

📅 May, 2020

I got accepted to the Archives-Unleashed datathon which was to be held at Montreal, QB, Canada. The event was canceled due to COVID-19.