



OLD DOMINION
UNIVERSITY

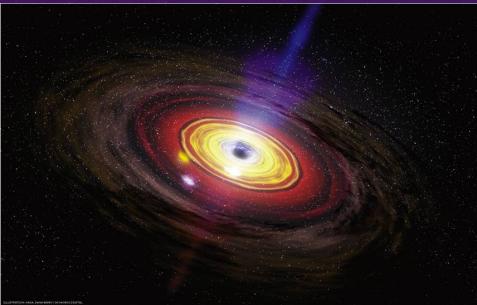
RESEARCH IN THE THE LAMP-SYS LAB

DR. JIAN WU

[HTTPS://WWW.CS.ODU.EDU/~JWU/](https://www.cs.odu.edu/~jwu/)

ABOUT ME

- PhD: 2004 - 2011
 - Astronomy & Astrophysics
 - Sloan Digital Sky Survey (SDSS)
 - Hubble Space Telescope (HST)
- Postdoc: 2011 – 2017, with Dr. C. Lee Giles at Penn State
 - CiteSeerX Tech Leader
- Assistant Teaching Professor: 2017 - 2018
- Assistant Professor (2018 –)
 - Web Science Digital Library Group (WS-DL)
 - Directed by Dr. Nelson and Dr. Weigle
 - Established the LAMPS-LAB



CURRENT MEMBERS

- Dr. Jian Wu – director
- 4 PhD students
 - Muntabir Choudhury (since Fall 2019)
 - Xin Wei (since Summer 2020)
 - Kehinde Ajayi (since Spring 2021)
 - Lamia Salsabil (since Spring 2021)
- 1 co-advised PhD student
 - Md Reshad Ul Hoque (since Spring 2019 with Dr. Jiang Li)
- Master's students
 - Martin Gryder (since Spring 2020)
 - Winston Shields (since Spring 2021)
- 1 Undergraduate students
 - Dominik Soos (since Spring 2022)



Muntabir Choudhury



Xin Wei



Reshad Hoque



Kehinde Ajayi



Lamia Salsabil



Winston Shields

RESEARCH DIRECTIONS

- Muntabir Choudhury: scientific document segmentation, focusing on electronic theses and dissertations
- Xin Wei: Semantic information extraction from content of scientific documents
- Kehinde Ajayi: Scientific data extraction and aggregation from scientific tables and figures
- Lamia Salsabil: Scientific document reproducibility analysis
- Md Reshad Ul Hoque: Claim verification and misinformation evidence retrieval
- Winston Shields: Web applications for scientific document and figure retrieval and annotations

PREVIOUS MEMBERS

- 1 PhD student
 - Richard Pates (since Spring 2020, on leave)
- 2 Master's students
 - Sami Uddin (Spring 2020 – Spring 2022, graduated)
 - Pei Wang (Fall 2019 – Summer 2020, transferred)
- 10+ Undergraduate students
 - Ryan Hiltabrand (Fall 2020—Fall 2021, ODU MS)
 - Gunnar Reiske (Summer 2019, Summer 2021, transferred)
 - Andrew Mayes (Spring 2019 – Fall 2019, graduated)
 - Bradley Dash (Fall 2018 – Spring 2019, graduated)



Gunnar Reiske



Andrew Mayes



Bradley Dash



Ryan Hiltabrand



Pei Wang



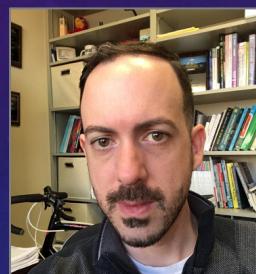
Sami Uddin

MISSION OF LAMP-SYS

Apply machine learning and deep learning techniques to build systems that solves multidisciplinary real-world problems using building blocks in natural language processing, scholarly big data, digital libraries, and information retrieval.

COLLABORATORS

- ODU
 - The WS-DL Group
 - Dr. Jiang Li (Electrical and Computer Engineering)
 - Dr. Jeremiah Still (Psychology)
- PSU
 - Dr. C. Lee Giles (Information Sciences and Technology)
 - Dr. Sarah Rajtmajer (Meta science)
 - Dr. Zi-kui Liu (Materials Science)
- Virginia Tech
 - Dr. Ed Fox (Computer Science)
 - Bill Ingram (University Libraries)
- Los Alamos National Laboratory
 - Dr. Diane Oyen (Information Sciences)
- Rochester Institute of Technology
 - Dr. Richard Zanibbi



SELECTED PROJECTS

- CiteSeerX: Developing a Sustainable Architecture to Support Scholarly Big Data
 - Collaborated with Penn State
- Mining Electronic Theses and Dissertations
 - Collaborated with Virginia Tech
- Synthetic Prediction markets with Algorithm Traders for Determining Experimental Reproducibility
 - Collaborated with Penn State
- Topological Relation-Based Image Analysis using Graphs
 - Collaborated with Los Alamos National Lab
- Fake scientific news detection
 - Collaborated with Dr. Jiang Li and Dr. Jeremiah Still at ODU



INFRASTRUCTURE

Usage	Hostname	#cores/ #threads	RAM	Storage	Type	Purchased
Computing	hubble	24/48	384GB	7TB	Physical	2019
Computing	hpc-highmem-3	48/96	322GB	2TB	Virtual	NA
Web	jwu-dev-mgmt	2/4	8GB	40GB	Virtual	2018
Web	penzias	--	--	--	Virtual	2020
Database	hawking	8	128GB	4TB	Virtual	2019
Database	peebles	10/20	128GB	6TB	Physical	2020
Database	chandra	10/20	128GB	6TB	Physical	2020
Storage	jansky	4/8	16GB	15TB	Physical	2019
Storage	giacconi	6/6	64GB	28TB	Physical	2020

* All storage is after RAID5

WHOM WE ARE LOOKING FOR

- Self-motivated
- Commitment
- Strong implementation
- Strong self-teaching ability
- Presentation

ADVISING PHILOSOPHY

- The beginning is the often hardest
- Be rigid and be flexible
- Short meetings
- Work in the lab
- Don't take it personally. Take it seriously!
- Be persistent

NOTABLE STUDENTS AND PUBLICATIONS

- **Lamia Salsabil**, Jian Wu, Muntabir Hasan Choudhury, William A. Ingram, Edward A. Fox, Sarah M. Rajtmajer, C. Lee Giles. "A Study of Computational Reproducibility using URLs Linking to Open Access Datasets and Software." In The Proceedings of the 2nd International Workshop on Scientific Knowledge. ([Sci-K 2021@WebConf](#)), April 26th, 2022.
- **Md Reshad UI Hoque**, Xin Wei, Muntabir Choudhury, Kehinde Ajayi, Martin Gryder, Jian Wu and Diane Oyen. "Segmenting Technical Drawing Figures in US Patents." In The Proceedings of the 2022 AAAI Workshop on Scientific Document Understanding. ([SDU 2022 @AAAI](#)), February 28 or March 1, 2022
- **Sami Uddin**, Bipasha Banerjee, Jian Wu, William A. Ingram, Edward A. Fox. "Building A Large Collection of Multi-domain Electronic Theses and Dissertations." In The Proceedings of the 2021 IEEE Big Data Conference. ([IEEE BigData 2021](#)), December 15–18, Virtual Event.
- Sarah Rajtmajer, Christopher Griffin, Jian Wu, Robert Fraleigh, Laxmaan Balaji, Anna Squicciarini, Anthony Kwasnica, David Pennock, Michael McLaughlin, Timothy Fritton, Nishanth Nakshatri, Arjun Menon, Sai Ajay Modukuri, Rajal Nivargi, **Xin Wei** and C. Lee Giles. "A synthetic prediction market for estimating confidence in published work." In Proceedings of the 36th AAAI Conference on Artificial Intelligence. ([AAAI 2022 Demo](#)), February 21–28, 2022, Vancouver, Canada
- **Muntabir Hasan Choudhury**, Himarsha R. Jayanetti, Jian Wu, William A. Ingram, Edward A. Fox. "Automatic Metadata Extraction Incorporating Visual Features from Scanned Electronic Theses and Dissertations." In the 2021 ACM/IEEE Joint Conference on Digital Libraries, ([JCDL 2021](#)). September 27–30, 2021.
- **Muntabir Hasan Choudhury**, Jian Wu, William A. Ingram, Edward A. Fox. "A Heuristic Baseline Method for Metadata Extraction from Scanned Electronic Theses and Dissertations." Poster in: Proceedings of the 2020 International Conference on Digital Libraries ([JCDL 2020](#)), August 1–5, 2020. Virtual Event. [[Best Poster Award](#)]
- Ming Gong, **Xin Wei**, Diane Oyen, Jian Wu, Martin Gryder, Liping Yang. "Recognizing Figure Labels in Patents." In: The 1st AAAI Workshop on Scientific Document Understanding @ AAAI 2021, ([SDU 2021](#)), Virtual Event.
- Jian Wu, Pei Wang, **Xin Wei**, Sarah Rajtmajer, C Lee Giles, and Christopher Griffin. "Accurate Acknowledgement Entity Recognition in CORD-19 Papers." In: Proceedings of the 1st Workshop on Scholarly Document Processing @ EMNLP 2020, ([SDP 2020](#)), November 19, 2020. Virtual Event
- **Md Reshad UI Hoque**, Jian Wu, Jiang Li, Chiman Kwan, Agnese Chiatti, and Dash Bradley. Searching for Evidence of Scientific News in Scholarly Big Data. In: *Proceedings of the 10th International Conference on Knowledge Capture (K-CAP 2019)*, November 19-21, 2019, Marina del Rey, CA, USA.
- Jian Wu, **Md Reshad UI Hoque**, Gunnar W. Reiske, Brenda T. Bradshaw, Holly D. Gaff, Jiang Li, Chiman Kwan. "A Comparative Study of Sequence Tagging Methods for Domain Knowledge Entity Recognition in Biomedical Papers." In: Proceedings of 2020 Joint Conference on Digital Libraries ([JCDL 2020](#)), August 1–5, 2020. Virtual Event
- Jian Wu, **Pei Wang**, **Xin Wei**, Sarah Rajtmajer, C Lee Giles, and Christopher Griffin. "Acknowledgement Entity Recognition in CORD-19 Papers." In: Proceedings of the 1st Workshop on Scholarly Document Processing @ EMNLP 2020, ([SDP 2020](#)), November 19, 2020.