

Laura W. Dozal

• Email: lwerthmann@arizona.edu • Website: lwdozal.github.io • LinkedIn: [./laurawerthmann/](https://www.linkedin.com/in/laurawerthmann/)

EDUCATION

PhD Candidate, iSchool, Current; Minor in Statistics and Data Science

Certificate in Computational Social Science

University of Arizona, Tucson, Arizona

Master of Arts in Communication, Culture, and Technology

Georgetown University, Washington, D.C. 2016

Bachelor of Arts in English Literature; Minor in Communication

St. Edward's University, Austin, Texas 2012 **Recognition: Dean's List, 2010**

EXPERTISE

Network Analysis • Machine Learning • Computational Social Science • Computer Vision • Data Mining • Information Diffusion • Predictive Modeling • Classification • Regression • Clustering • Hypothesis Testing • Bayesian Analysis • Gaussian Models • Recursive and Convolutional Neural Networks • Dimensionality Reduction • Multiclass Classification • Multilevel Modeling • Natural Language Processing • Word Embeddings • Visual Frames • Social Movements •

Programming Languages / Tools: R - tidyverse, ggplot2, Survival, bnlearn • Python - NumPy, pandas, matplotlib, scikit-learn, py-torch, spaCy, TensorFlow, keras • Jupyter, Spyder • SQL • OpenCV • LaTeX • Git/Github • Javascript, D3.js, HTML • ArcGIS • Tableau • Anaconda • InDesign • Open Refine • UX analysis and research

Bilingual: Fluent in English and Spanish Languages

SPEAKING ENGAGEMENTS and CONFERENCES

- **Winter Workshop on Complex Systems (Santa Fe Institute)**, 2024 Presenter and Participant
 - **Bibliometric Data: Publishing Behavior of University of Rochester Researchers**, University of Rochester Libraries presenter for LEADING Data Fellows
 - **University of Arizona iSchool Colloquium**, *Working with Foundational Open-Source Science Skills (FOSS) for Dissertation Data and Reproducibility*
 - **Narrative Frames in Topical Images: Mixed Methods Approach to Understanding Social Movements**, ASIS&T Workshop Panelist: "Turning Social Informatics Research into Action in a Changing Moment", London UK, Oct. 2023
 - **Posit::Conf2023**, Opportunity Scholar, September 2023
 - **Doctoral Colloquium**, iConference, 2022
-

PUBLICATIONS

- **Visual and Textual Analysis of the Online Anti-Femicide Movement in Mexico: A Narrative Network Approach**, Conference of Complexity Science 2024 Poster Acceptance; Exeter, England
- (Accepted, not published), **Visual and Textual Analysis of the Online Anti-Femicide Movement in Mexico: A Narrative Network Approach**, ASONAM2024, IEEE Conference; and ACM GOODIT2024 Conference
- **Visual and Textual Narrative Analysis of the Online Anti-Femicide Movement in Mexico: A Narrative Network Approach**, Association for Computational Linguistics (ACL) Conference - WOAHP workshop poster, 2024
- **Narrative Frames in Topical Images: Mixed Methods Approach to Understanding Social Movements**, ASIS&T Workshop Panelist: "Turning Social Informatics Research into Action in a Changing Moment", London UK, October 2023
- **Mixed Methods Framework for Understanding Visual Frames in Social Movements**, ASIS&T Conference, Long Paper Publication, London UK, October 2023
- **Mixed Methods for Understanding Visual Frames in Social Movements**, International Conference on Computational Social Science (IS2C2), Poster, Chicago, July 2022
- **Extracting Space Situational Awareness Events from News Text**, LREC Conference. Paper Publication May 2022
- **Street Art as Visual Information: Mixed Methods Approach to Analyzing Community Spaces**, ASIS&T Poster, 11/2021
- **Detecting the Spread of Propaganda Bias Online: Comparing HAN and BERT Models**, TPRC Conference, Paper Publication, 4/15/2021
- **Data Visualization - Old North Pilgrimage Project**, Permanent Installation, Georgetown University, 4/21/2016
- **Network Installations - An Art Technology**, Paper Presenter, PCA/ACA 06/09/2015, STGlobal Conference, DC 11/2015
- **Street Art Dialogue**, Gnovis Journal, Dec. 18, 2014 (Peer Reviewed Publication)

EXPERIENCE & RESEARCH**Data Science Experience**

Instructor, ISTA 116: Statistical Foundations for the Information Age | College of Information, University of Arizona, 08/2024 – Present

- Instructor on record for undergraduate course that covers understanding uncertainty and variation in modern data: data summarization and description, rules of counting and basic probability, data visualization, graphical data summaries, working with large data sets, prediction of stochastic outputs from quantitative inputs. Operations with statistical computer packages such as R.

Data Science Fellow | Laboratory of Applied Science - Summer Conference of Applied Data Science; North Carolina State University, Raleigh, North Carolina, 06/2024 – 07/2024

- NLP researcher comparing chunk-size methods for text summarization. Used extractive summarization and abstractive summarization, Retrieval Augmentation Generation (RAG) to test different text segmentation methods. Used and created vector embedding database for evaluation using RAGAs, an LLM RAG evaluator. Used the 2024 TREC Challenge data-MSMarco docs and QA for analysis.
- Implemented two video object tracking models: zero-shot transformers - OWL-ViT and LMM gpt-4o on noisy data- dashcam footage of a ride in Afghanistan provided by LAS staff. The goal was to explore improvements for tracking object obfuscation and detecting unknown/unconventional objects in the noisy surveillance video data. We found that pre-trained image models provided the most accurate structured data outputs.

Complexity Science Researcher | Winter Workshop on Complex Systems (Santa Fe Institute), Spain, 01/2024

- Ongoing research of spreading infections (SI) phenomenon through honeybee food transmission and communication (trophallaxis) using complex systems methods including autocorrelation networks, temporal networks, burstiness, and social role clustering.

R4R (Roots for Resilience) Dissertation Fellowship, 2023 Cohort | *University of Arizona Data Science Institute, Tucson, AZ, 08/2023 – 01/2024*

- *Data Management*: Created and maintained a Data Management Plan using Open Science Framework in collaboration with CyVerse. Use of DataOps and pipelines, Version Control - github, and Reproducibility- Docker.
- *Dissertation*: Map Building and Image Analysis for social movement group in Mexico [Elastienennombre.org](https://elastienennombre.org). Maintain multimodal dataset in MySQL, create interactive Leaflet map. Used Github to apply computer vision methods including localized feature matching, segmentation, homography, recognition and Bayesian modeling.

LEADING Data Science Fellowship 2023 Cohort | *University of Rochester Remote 06/2023 – 01/2024*

- Lead creation of an Open Source Data Pipeline to collect and analyze bibliometric data from University of Rochester Researchers and identify publication behavior. Collected 500K+ metadata using APIs from various repositories, cleaned and processed the data to set-up for analysis and better reproducibility.
- Used trend analysis and network analysis methods including regression (decision tree), time series analysis, structural analysis, and prediction analysis for network validation of researcher behavior.

Undergraduate Instructor | *University of Arizona School of Information Tucson, AZ, 08/2019 – 08/2024*

- Instructor for ESOC 130 – Computational Thinking and Doing. Introduction to python topics including functions, lists, classes, and logic. Provide computational theory for problem solving with algorithms and other computational models.
- Instructor for ESOC 211 – Collaborating in Online Communities. The class focuses on what humans do, how they present themselves, and how they do the work of collaboration in online contexts by considering how people create a sense of community, maintain group connections, and cooperate with one another to bring about a particular outcome. Students learn online collaboration skills, and alternative thinking processes on how to approach online communities. introduced digital tools such as Lucid Chart, Adobe Spark, and digital content methods such as user personas, metadata, and SEO. They also learn genre analysis and visual rhetoric.
- Instructor for ESOC 300 – Digital Storytelling and Culture. The class focuses on how story, audience engagement, narrative structure, and digital mediums have been harnessed to understand how stories are essential to our experience. Students complete projects with elements of community, identity, memory, and perspective.

EXPERIENCE & RESEARCH (Data Science Experience Continued ...)**Data Science Grad Assistant | University Analytics & Institutional Research, UA, Tucson, 08/2021 – 06/2023**

- *NLP Sentiment Analysis*: Analyze 4M+ student course survey responses using Natural language processing to identify sentiment, word frequencies and word synonyms for curriculum in various departments on campus. Increase faculty participation in curriculum review across schools on campus.
- *Machine Learning - Churn Analysis*: Spearheaded the Women in Science and Engineering Retention report project which uses survival and churn analysis. These analyses review churn or dropout rates of 37K+ students over ten years. The analysis incorporates machine learning applications including decision tree and random forest models along with other time-based statistical methods. I Identified student financial, cultural, academic, and demographic characteristics that either debilitate or foster success towards graduating college. Increased awareness across university departments to enable discussion and student experience development.
- *Dashboard Creation - Analytical Problem solving*: Using automated and qualitative approaches, I developed End-to-end application of tuition and headcount dashboard visualizations and analysis using R and SQL. Create aggregated data tables from an internal Oracle data warehouse that include descriptive statistics and inflation information of campus-wide Net Tuition Revenue (NTR). The dashboard is used by leaders and policy makers on campus to make informed decisions.

Computational Tools and Thinking - Teaching Fellow | University of Texas Austin, 06/2022 - 7/2022

- Developed ten, two-hour lesson plans for a computational tools and thinking survey course that covered R, Python, Social Network Analysis, Design Theory, and coding best practices. Taught to a wide range of experience levels that included 25 high achieving undergraduate students from STEM and Social Sciences.
- Held office hours to support outside research and enabled the development of a 20-month research design process. Participated in community building activities with other distinguished fellows as well as with our students. Maintained mentor level contact with students to aid in coding and analysis applications for research.

Data Science Ambassador | University Arizona Data Science Institute, Tucson 08/2021 – 08/2022

- Data Carpentries Instructor as of 2/14/2022 -Live coding instructor and helper for Data Analysis and Visualization with R for Social Scientists, Introduction to Python, Open Refine, and other data science skills.
- Data Science community support and resource for the CSS certificate and iSchool. Created and Organized Projects in Progress. Fostered a Data Science community environment for researchers on campus to collaborate and share projects ideas and progress. Helped consult 3 - 5 projects over the span of two semester (over 20 researchers) using data science methods monthly.
- Helped organize Stanford-based Women in Data Science satellite conference and led the Computational Social Science panel as an organizer and moderator. April 2022 event had around 400 participants on and offline.

Data Analyst Research Assistant | Arizona Medical Education Research Initiative, UA, Tucson, 01/2021 – 08/2021

- *Multivariate Statistical Analysis* Applied statistical analysis an inference to survey and collected data to review and test the University of Arizona's Medical School's performance for future accreditation. Types of data include survey data, archived data, and collected government data from large data systems. Used both R and Python to apply analysis of variance, chi-square tests, validation checks, regression, and correlation analysis.
- *Program Use*: Worked with Power BI, Excel, R and Python to clean, manage, transform, analyze, iterate, and visualize data. Created reports from various sized data sets using statistical analysis and visualization.

Research Analyst | Security Analysis Co-Lab. University of Arizona, Tucson, AZ, 03/2020 – 9/2021

- Lead data management and repository infrastructure foundation for survey data from every state in El Salvador. Prepared and analyzed a randomized control treatment to schools in high crime areas. Applied simple statistics, hypothesis testing, and regression analysis to collected data of large-scale survey instrument roll-out.
- Assisted with grant proposal writing, coordinated with government and community collaborators to identify research parameters and methodology. Used a logical framework and theory of change to compose project.

EXPERIENCE & RESEARCH (Data Science Experience Continued ...)**Survey Data Manager | Minerva Project: University of Arizona, Tucson, AZ, Summer 2021**

- NSF Grant funded project: Contributed to maintenance of a pipeline for 8k+ unique and personalized surveys in English, Spanish, Pashto and Dari. Applied data cleaning, wrangling, and statistical analysis to a databank of crowd-sourced history of armed conflict in Afghanistan and Colombia for the US Department of Defense.
- Troubleshooted and applied forensic analysis to problem solve issues in the pipeline and iterate over processes for optimization. Assessed project state and automated forecasting of coverage to establish future project needs. Completed Time Series visualizations for dynamic datasets by reformatting and manipulating data in R to show the territorial control by Criminal Organizations.

Research Projects

Dissertation Research: Image Analysis, Fall 2022 - Present Focused on mixed methods approach towards identifying narratives through topical images using *network analysis* (Hierarchical Graphs) and *computer vision* (Feature Recognition, CNN, Classification). Qualitative methods implement a collaborative and interpretive discussion on my results with a focus group of domain experts to solidify findings. Ultimately the goal is to create a knowledge graph of narrative relations to identify overall implications of image messaging for a social movement.

Data Mining: Collected thousands of social movement images and metadata including comments, likes, hashtags, URLs, and automated descriptions from Instagram using selenium and beautiful soup to pull content and then implement network science to classify narrative frames using deep model feature learning and sentiment analysis.

Knowledge Graph/Ontology Research: Final Project, 2022 Reviewed knowledge graph and Ontology methods for visual storytelling, semantic attention, and image classification for final paper. Particularly, methods of scene graph applications, which represent relational knowledge that can be modeled with symbolic data from texts or knowledge graphs. Explored Knowledge graph use of the semantic web was explored to understand image processing. Also included hierarchical cross-model format with focus on video and text retrieval entailing methods hierarchical models, temporal segmentation algorithm, bidirectional recursive neural networks (Bi-RNN), cosine similarity as a loss function, and self-attention features.

TPRC Conference, Publication, 4/15/2021 *Deep Models*: With partner, compared two natural language processing models on a dataset composed of labeled propaganda data. We reviewed off the shelf BERT model and the Hierarchical Attention Network (HAN) model and found they both provide different accuracy levels, with BERT maintaining better results. The main fault of the HAN model is that it was not able to account for the unbalanced aspect of our data which was labeled in a binary (propaganda - 1, not propaganda-0).

LEADERSHIP

AI Steering Committee, University of Arizona, 2023 - Present

PhD Community Manager (University of Arizona iSchool) 2021 - Present

Data Carpentries Instructor – Lead coding workshops in R, Python, Open Refine, and data science skills 2022 –Present

NLP Reading Group UofA, 2019-Present

ASIS&T SIG – Visualization, Sound, and Image: Secretary Chair 2021-2022

WiDS-Tucson (Women in Data Science), Panel Sessions Organizer, Data Blitz Committees 2021 -2022

Data Science Ambassador- Data Science liaison for the CSS certificate and iSchool, Datathon Organizer 2021 – 2022

RezBaz - Presenter & Moderator, Network Analysis 2021; R Visualization 2021 and 2022

iSchool Colloquium, Organizer: Moderated Lectures, led school-wide Faculty Research Blitz 2020 – 2021

COURSEWORK AND CERTIFICATIONS

Advanced Natural Language Processing, CS – Fall 2024

Introduction to Computer Vision, Computer Science, Fall 2023

Theory of Statistics School of Statistics, Spring 2023

Advanced Data visualization - Computer Science – Spring 2023

COURSEWORK AND CERTIFICATIONS (Continued...)

Text Analysis Pedagogy Institute – JSTOR & University of Arizona, Summer 2022

Introduction to Artificial Intelligence– School of Information, Spring 2022

Bias in Information Retrieval Systems (Graduate Seminar) – School of Information, Spring 2022

Probabilistic Graphical Models, Computer Science – Spring 2022

Theory of Probability – School of Statistics, Fall 2021

Visual Epistemologies of Media and Technology (Graduate Seminar) - School of Information, Spring 2021

Applied Social Network Analysis– School of Information, Spring 2021

Social Movements and Collective Action – School of Sociology, Fall 2020

Neural Networks– School of Information, Fall 2020

Bayesian Models and Inference– School of Information, Fall 2020

Human Rights and Technology – School of Information, Fall 2020

Intro to Machine Learning- Electrical and Computer Engineering, Spring 2020

Micro-dynamics of Political and Organized Criminal Violence – Government and Public Policy, Spring 2020

Science, Health, Engineering Policy, and Diplomacy; Sustainable Development for the Americas Conference, Fall 2019

Presenting Data and Information- Edward Tufte, 2016

PROFESSIONAL EXPERIENCE

Community Outreach and Program Coordinator – *Lutheran Social Services of the Southwest K-12 Refugee Education (AmeriCorps State Member), Tucson, AZ, 08/2016 – 08/2019*

- Arizona Serve State Member working with the non-profit organization, LSS – K-12 Refugee Education CENTER, by providing academic support, empowerment, and leadership skills to refugee students in Tucson, Arizona. Provided assessment and outreach for Tucson schools with large refugee student populations by developing a database to track and identify community resources and partnerships. Connected students to local academic advancement resources.
- Created a sustainable measurement process to identify the impact of LSS-SW Education services based on student survey and school-based data. Produced a report of the year's work, updated Refugee Youth Services Handbook, and share with stakeholders. Kept LSS-SW a primary education resource for Tucson refugees.
- Helped students with website creation and create lesson plans and workshops to help students get better acclimated to using digital files and performing research on the internet. Used web design best-practices and maintained a website showcasing student work that expresses themselves and their stories of refugee immigration. Review communication apps for students and teachers.

Digital Content Specialist – *Tetakawi, Tucson, AZ, 08/2016 – 05/2019*

- Conducted field research on the Mexican labor market for advisory services branch. Provide report editing, and data check to research findings. Researched industry resources on economic, regulatory and policy events in Mexico and U.S. trade relations on company website.
- Worked with SEO, Google analytics and CMS platforms for web optimization. Create content using design and basic principles for blogs, e-books, white-papers, infographics, graphics, and market reports. UX research and design for website re-branding. Including wire-framing, prototyping, testing, surveying, and content creation.

Web Designer / Content Management - *McCourt School of Public Policy, Georgetown University, Washington, D.C. Summer 2016*

- Applied design thinking consulting to help the marketing, IT, and communications director transition to new a website platform. Solved Problems with coding elements using bootstrap and Drupal platforms. Applied UX design methodology and analysis. <https://mccourt.georgetown.edu/>. Applied quality assurance review and web design best practices for multiple website projects.

PROFESSIONAL EXPERIENCE (Continued...)**Teaching / Research Assistant - *Georgetown University, Washington, D.C.* 08/2015 – 07/2016**

- **Fundamentals of Technology:** Worked with a team of TAs and professors to implement syllabus and in- class participation. Led breakout discussion sessions to guide students with leadership and interpersonal skills as they develop an understanding of the material. Maintained a strong understanding of course content to answer questions on technology assessment and guide students on assignments, project development and exam instructions.
- **Content Analysis Research:** Applied data management skills to review data sets and reports. Developed analytics guide to further the study. Used the predictive analytics software, SPSS, to clean and create multiple databases to observe content and survey records. Applied critical thinking and analysis to data. Created annotated bibliographies, and edited reports.

Digital Experience Intern – *Archives of American Art, Smithsonian Institution, Washington, D.C.* Summer 2015

- Implemented website development, quality assurance, design research for logo redesign. Audited and managed website content for transition to the Drupal platform. Coordinated with staff for Drupal trainings and social media initiatives. Researched website best practices and current art news. Coordinated social media strategies with team. Created SEO metadata. Awarded the Horowitz-Fraud Minority Scholar.
<https://www.aaa.si.edu/>