MONTANA STATE UNIVERSITY | LIBRARY

Organization: Montana State University Library
Primary mentors: Jason A. Clark, Lead for Research Optimization, Analytics, and Data Services (ROADS) at Montana State University Library; Hannah McKelvey, Electronic Resources & Discovery Services Library at Montana State University Library

Project Title(s)	Recognizing Bias in Collection Metrics: Analyzing Institutional Scholarship and Grant Networks to Inform Collection Development Decisions
Description	Often, decisions about collection development are made by looking at usage of resources (for example, how many times a book chapter or article was downloaded). While this is a typical approach for assessing library collections, these metrics can be biased — subject areas with more students and faculty are logically going to have more usage than their smaller counterparts. However, by analyzing institutional scholarship such as publications and grant award data, bias in collection assessment strategies can be recognized and considered in the decision making process. This project will analyze Montana State University (MSU) scholarship published from 1960-current, using a corpus of research generated via OpenAlex. In addition to this dataset, fellows will also have access to MSU grant award data provided and eResource subscription data. Together, these 3 datasets can be used to create data visualizations, such as collection profiles or researcher resource recommendations, in order to tell a compelling story about how resources might be used over time at the MSU Library. These types of visualizations would complement numerically-based visualizations. Fellows will have the opportunity to create deliverables that could include: ● Visualizations of grant and scholarship networks in place at MSU ● Personal collection profiles that can be shared with individual researchers ● Using the aggregated data to power a "Collection AIssistant" (An AI-powered conversational agent)
Problems/ Research Questions	Fellows will be encouraged to develop complementary research questions and solutions. Initial research questions include: ● Are these data snapshots enough to make informed collection development decisions? ● What additional metrics and analytics might be needed to reach the full impact of this project? ● How can these aggregated data sources be used for novel visualizations or research recommendations?

Techniques	• Web scraping and text mining • Network Analysis • Natural language processing • Data Visualization
Tools/ Languages used	The project team has expertise in Python, Javascript, SQL, and Excel. Some capacity exists to support fellows interested in network analysis and NoSQL/graph databases.
Data	**Description:** Montana State University grant award data provided by local offices or national aggregators, OpenAlex research corpus with MSU authors, and MSU eResources subscriptions list. MSU Awarded Grant data (1982-2021) MSU Scholarship (OpenAlex) - showing JSON from API; have data as SQL as well. • MSU: https://explore.openalex.org/institutions/I23732399 • MSU total works: https://api.openalex.org/works?filter=institutions.id:I23732399 • MSU total works 2022: https://api.openalex.org/works?filter=institutions.id:I23732399,publication_year:2022 MSU List of Journals (subscriptions) **Data Type:** JSON, CSV, SQL **Data Size:** MBs
Outcome	Published papers, presentations, datasets, software code.
Milestone Timeline	Six months per selected project
References	Collection Assessment Dashboard at the MSU Library: https://montana.libinsight.com/collection-assessment Linked Biosciences at Montana State University (prototype) https://github.com/jasonclark/linked-people-bioscience https://arc.lib.montana.edu/linked-bioscience/graph/expertise-bubble.html