Organization: Schoenberg Institute for Manuscript Studies
Primary mentor: Lynn Ransom
Secondary mentor(s): Doug Emery, L. P. Coladangelo

. Project Title	Digging into Digital Scriptorium 2.0: Data Analysis in an LOD Environment
Description	Digital Scriptorium (DS) is a growing consortium of American cultural heritage institutions committed to offering open online access to their collections of premodern manuscripts. The DS consortium comprises a variety of institutions, from large research universities to small public libraries and private museums that contribute metadata toward this end. DS 2.0 is the recently reimagined and rebuilt online platform providing an inclusive, open access online union catalog for all premodern manuscripts in the United States. Using Wikibase to collect and aggregate member institutions' metadata, DS 2.0 offers a unique service for finding and researching manuscripts across the US that relies on structured metadata linked to external authorities and LOD datasets, including Wikidata, thus connecting DS data to a vast network of knowledge graphs. Using DS 2.0 as a case study, the goal of this project is to provide an opportunity for a LEADING fellow to investigate the benefits and challenges of LOD in a humanities research context. Specific tasks will involve investigating how Wikibase data can be used in conjunction with other LOD datasets and sources like Wikidata, the Getty Vocabularies, and the Virtual International Authority File, among others.

Problems/ Research Questions	The following are suggested research questions. The fellow can develop others in consultation with the project team. 1) What are research implications for creating and using LOD? 2) What data is meaningful in an LOD environment? What makes it meaningful? 3) What issues arise with DS data relating to cultural heritage representation in an LOD environment? 4) What patterns related to the study of premodern manuscripts emerge from an analysis of DS data.
Techniques	1. LOD 2. SPARQL service query creation and execution 3. Data Analysis 4. Other relevant techniques as identified in consultation with project team
Tools/ Languages used	SPARQL, Wikidata, RDF, Python (as needed)
Data	The fellow will have access to the DS Wikibase and SPARQL query service upon acceptance. Description: Metadata related to the production and transmission of global premodern manuscripts held in US collections Data Type: Wikibase items, entities, and properties Data Size: ca. 5000+ records to begin with
Outcome	1) Testing and analysis of DS data quality, especially in regard to interactions with other LOD datasets 2) SPARQL service queries for publication on DS website using DS data with other LOD datasets 3) Peer-reviewed article summarizing and assessing results related to research questions for publications
Milestone Timeline	Month 1-2: Gaining familiarity with dataset and developing basic queries for testing and analysis. Month 3-4: Developing advanced SPARQL queries, including but not limited to service queries

	Months 5-6: Analysis and assessment; article preparation
References	https://digital-scriptorium.org/