**Project Title**: Programmatic Evaluation of Subjects in The Portal to Texas History

**Description**
The Portal to Texas History currently hosts over 1.4 million items, including handwritten and printed text, images, audio-visual items, and other cultural heritage materials from across the state of Texas. All items in the Portal have the same UNTL metadata format which requires a minimum number of subject values, including at least one hierarchical term from our locally-managed controlled subject vocabulary. This project would attempt to improve methods of evaluating subject usage as well as possible development of tools to assist editors with subject assignment in metadata records.

**Problems/Research Questions**
- What is the distribution of UNTL-BS terms (and other subject types) across Portal records?
- How are various terms co-located?
- Can the appropriateness of subject values be determined programmatically?
- Can this evaluation (or machine learning) provide suggestions to editors for assigning subjects?

**Techniques**
Data Mining, Visualization, Network Analysis

**Tools/Languages used**
Python, R, Tableau, GitHub
| Data | Description: This project will make use of aggregated metadata records for over 1.4 million items in The Portal to Texas History  
Data Type: XML files, JSON files, Markdown |
| --- | --- |
| Outcome | • Develop a framework for quantitatively and qualitatively evaluate subject usage programatically.  
• Identify ways that subject evaluation could improve tools for editors assigning subjects during metadata creation. |
| Milestone Timeline | • Identify subject usage in the Portal across resource types and collections  
• Create visualization of existing collection data  
• Develop criteria for using programmatic means to evaluate or suggest terms |
| References | The Portal to Texas History – [https://texashistory.unt.edu/](https://texashistory.unt.edu/)  
University of North Texas Libraries Browse Subjects (UNTL-BS) – [https://digital2.library.unt.edu/subjects/browse/](https://digital2.library.unt.edu/subjects/browse/) |