KATHERINE HART WEIMER
Kelley Center for Government Information, Data, and Geospatial Services
Fondren Library
Rice University
Email: khw2@rice.edu

Katherine (Kathy) Hart Weimer is head of the Kelley Center for Government Information, Data and Geospatial Services at Rice University’s Fondren Library. Kathy has worked with maps, geographic information, library catalogs and metadata throughout her library career, and has fostered awareness of the geographic nature of library collections through building map-based search interfaces. Her interests include the GeoHumanities and related topics of spatial thinking, geo-information literacy, and gazetteer development. Weimer is co-editor of the Journal of Map & Geography Libraries and past Chair of ALA’s Map and Geospatial Information Round Table. She co-founded, and is co-chair of the GeoHumanities Special Interest Group of the Alliance of Digital Humanities Organizations (ADHO) where, among other things, she has facilitated the planning of workshops, most recently on linked open data and gazetteers. Among Kathy’ research is serving as project lead for Mapping Historic Aggieland (it’s direct link). Her current research is consulting with the Mellon funded Linked Data for Libraries (LD4L) project’s Cartographic Extension Sub-group to extend library catalog data into a linked environment, specifically for maps and cartographic resources.

Position Statement

Geographic content permeates library collections, whether through places described in books, depicted on maps, newspaper articles or photographs, as well as the location of authors and publishers. Spatially mediated discovery, through a single-point access is quite complex, and unfortunately may supply haphazard results due to variations of content standards across different schemas. For example, library catalog data is quite structured, and due to historic content standards, is often a simplified, aggregate representation of intellectual (geographic) content. This is a very different approach than what is found in metadata found in library digital repositories (i.e. Dublin Core), which often contains unstructured metadata with varying levels of detail. Different still is the robust and precise geospatial metadata found within a GIS. How can we integrate these various worlds into a single search if the underlying content standards are based on different assumptions? Additionally, consideration must be given to known geographic issues of scale, spatial footprints, and temporality. Integration in a single search must take into account the variation of input standards and assume messy results.

One approach to expand the catalog data and therefore, the discoverability of research objects with place references is with linked open data. Through LOD, scholarship is facilitated both within and more importantly, outside the library environment. Library catalog place name data is evolving to LOD. The long-term desired outcome of the LD4L project is to facilitate this flow of information to the broader scholarly environment. There are numerous scholarly gazetteers with detailed curated place name references, among them, DINAA (Digital Index of North American Archaeology),
Pleiades, and Pelagios among many others. Further, there are broad based LOD resources, including the Getty Thesaurus of Geographic Names, GeoNames and Wikipedia/DBPedia, which provide robust gazetteer data. Challenges, here include the complication that that each discipline will approach the definition of a place, a region, and place-related time periods from their particular disciplinary lenses. Ultimately, however, the integration of library (spatial) search to a wide variety of linked resources will only serve to expand the scholarship of place.

Another approach to the challenges of single search may be by using a combination of spatial search (using geographic coordinates) and natural language processing of the textual elements inside gazetteers. NLP could be employed to evaluate and process gazetteer data (including library catalog ‘Name Authority’ data and other digital gazetteers). From a search request, a NLP/spatial tool could evaluate dates, alternate names, spatial extent, etc. from the gazetteers, and then perform a more precise spatial search, therefore improving search results.

Scholarship will greatly benefit as we continue to pursue these spatial discovery questions. The ultimate goal of a modern and relevant library discovery system should be a seamlessly integration with the wider semantic web, therefore, intellectual access to library collections alongside other GLAM and scholarly resources. It is part of human nature to want to understand the why of “where.” Surfacing human knowledge through spatial search will allow for more robust ‘where’ questions, and ultimately more meaningful answers.