SARA LAFIA
Department of Geography
University of California, Santa Barbara
Email: sara.lafia@gmail.com

Sara Lafia is a graduate student in the Geography Department at UCSB. She works in partnership with the UCSB library and the Center for Spatial Studies on improving the spatial discovery of research data and documents. Her research addresses the question of how to spatially enable discovery of connected data and publications in a setting that allows for mapping and analysis using a Geographic Information System. She is also interested in the application of spatialization frameworks to non-spatial data, such as text, to gain new insights about themes of contents across formats. Her background is in Urban and Regional Planning. She has experience working on projects in collaboration with the Region 10 Environmental Planning Agency, the South Coast Air Quality Management District, and the Jet Propulsion Laboratory, applying GIS to assess social issues while also improving the accessibility of spatial tools.

Challenges

Location is an integrator of disciplinary perspectives and enables data producers to increase the discoverability of their products by offering a framework upon which data from multiple sources can be integrated. Platforms, for example Esri Open Data, have enabled format-agnostic approaches for data hosting and description at UCSB, which allows for a single-point of access across already hosted services, types, and domains. There are great opportunities for coordinating data production, dissemination, and use across university campuses specifically. However, while abundant in principle, data portals that span multiple repositories and integrate contents based on location are still far from common in practice. Current efforts, like Geoblacklight, tend to focus specifically on the discovery of more traditional cartographic products, such as geospatial data, while other types of data, such as interviews or historic texts, fall by the wayside. How can we facilitate the comparable spatial discovery of qualitative data, such as ethnographic surveys and text, which are not described in spatially explicit ways?

My research has produced a simple and extensible workflow for describing research objects as a library service and for spatially integrating datasets across repositories. This addresses current common practices for data publishing in academia, which often results in datasets that are not easily discovered, are hard to integrate across domains, and typically are not linked to publications about them. For example, archaeological observations and specimen data collections that share a spatial extent in Mesoamerica, should be concurrently discoverable. Relevant publications or other documents that reference the datasets should also be linked. Additionally, the notion of integrating across scales poses challenges. How should the location of data observations be described across multiple spatial scales?
Lastly, expanding the notion of spatial discovery to topic spaces offers another view of data that is non-geographic. For example, recent efforts to produce a self-organizing map of campus research, based on the subject descriptions of dissertations, yielded interesting insights into clusters of researcher themes. Coupling topic modeling with place-based views of research, obtained from geoparsing the same research descriptions, allows for new insights, offering two entry points for discovery through both geographic and topic spaces. How can the application of techniques, such as topic modeling, to unstructured text communicate collaboration gaps and motivate further research across disciplines?