A Proposal for an Empirical Humanities and Creative Social Sciences Research Infrastructure Initiative

Lindsay Poirier

Science and Technology Studies, Rensselaer Polytechnic Institute

ABSTRACT

In this poster, I present several proposals for support services that can help advance adoption of RDA recommendations in the empirical humanities. I summarize ethnographic research conducted in the first 6 months of my RDA Data Share Fellowship on the barriers to adopting data sharing practices in the empirical humanities. I then go on to outline a proposal for an Empirical Humanities and Creative Social Sciences Research Infrastructure Initiative (EHCSSRII) that aims to support 1) organizational infrastructure, 2) technical data sharing services and infrastructure, and 3) critical data literacy education. I discuss early efforts to advance this initiative, along with some of the early challenges we've run into.
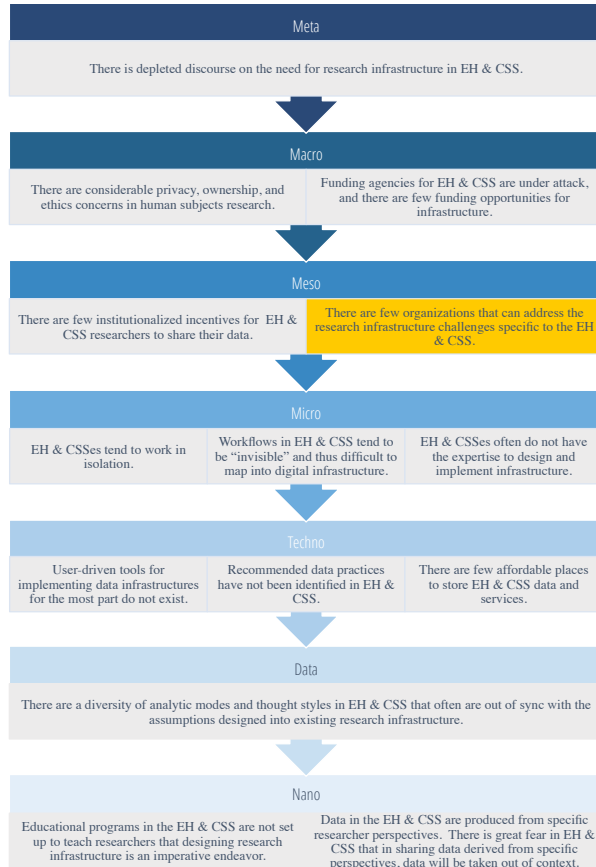
RELEVANT RESEARCH DOMAINS

Collaborative Hermeneutics
- Researchers address (continually iterating) sets of shared questions
- Researchers also toggle between their own interpretation and those of other researchers focused on "the same" data
- Researchers recognize that interpretation is not "subjective" but always informed by situated perspectives

Empirical Humanities & Creative Social Sciences
- Include history, folklore, cultural anthropology, and related fields
- Researchers *produce* data for cultural analysis, through interviews and other methods informed by diverse analytic traditions
- Often not hypothesis-driven
- Aiming to characterize both common themes and obscurities
- Researchers not trying to *reproduce* data but instead to layer theoretically informed thick descriptions of data

SPECIFIC DATA SHARING CHALLENGES IN THE EH & CSS

Meta

There is depleted discourse on the need for research infrastructure in EH & CSS.

Macro

There are considerable privacy, ownership, and ethics concerns in human subjects research.

Funding agencies for EH & CSS are under attack, and there are few funding opportunities for infrastructure.

Meso

There are few institutionalized incentives for EH & CSS researchers to share their data.

There are few organizations that can address the research infrastructure challenges specific to the EH & CSS.

Micro

EH & CSSes tend to work in isolation.

Workflows in EH & CSS tend to be "invisible" and thus difficult to map into digital infrastructure.

EH & CSSes often do not have the expertise to design and implement infrastructure.

Techno

User-driven tools for implementing data infrastructures for the most part do not exist.

Recommended data practices have not been identified in EH & CSS.

There are few affordable places to store EH & CSS data and services.

Data

There are a diversity of analytic modes and thought styles in EH & CSS that often are out of sync with the assumptions designed into existing research infrastructure.

Nano

Educational programs in the EH & CSS are not set up to teach researchers that designing research infrastructure is an imperative endeavor.

Data in the EH & CSS are produced from specific researcher perspectives. There is great fear in EH & CSS that in sharing data derived from specific perspectives, data will be taken out of context.
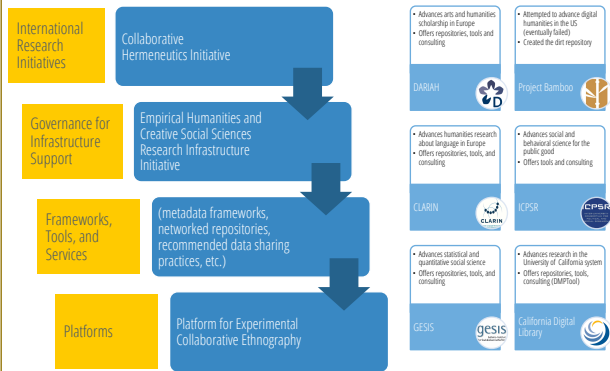
COLLABORATIVE HERMENEUTICS INITIATIVE (CHI)

CHI aims to interweave hermeneutic traditions in the humanities and social sciences with research and findings in scientific domains such as the life sciences, engineering, computer science, and medicine in order to answer complex questions about contemporary global problems. In bringing diverse, geographically-dispersed and interdisciplinary perspectives to the characterization of data, CHI seeks to advance what we've termed "collaborative hermeneutics." The initiative both seeks to advance understanding of contemporary problems, while also theorizing the research infrastructure needed to make possible the integration of data and analyses across scales, disciplines, and cultures.
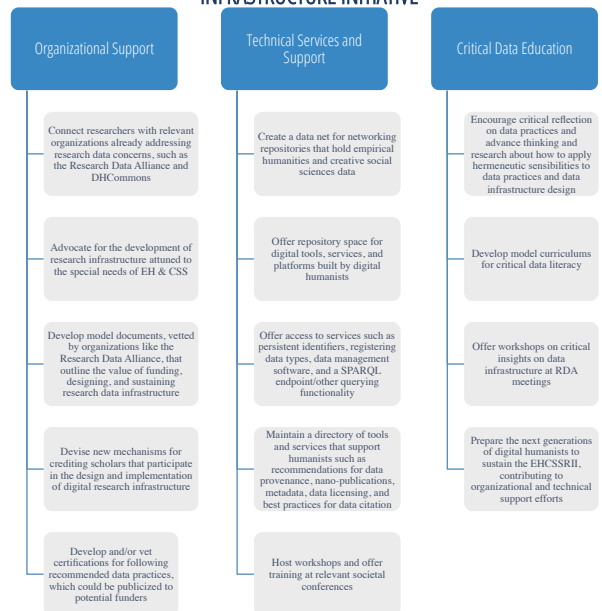
PROPOSAL FOR AN ECOLOGY OF RESEARCH INFRASTRUCTURE SUPPORT

We suggest that in order to advance an international collaborative hermeneutics initiative, we need governance for research infrastructure support that is attuned to the particular aims and challenges of the empirical humanities and creative social sciences. We hope that such support can network and vet existing research frameworks, tools, and services according to domain-specific challenges, as well as propose new strategies when existing frameworks, tools, and services do not align with domain-specific commitments.

International Research Initiatives — Collaborative Hermeneutics Initiative

Governance for Infrastructure Support — Empirical Humanities and Creative Social Sciences Research Infrastructure Initiative

Frameworks, Tools, and Services — (metadata frameworks, networked repositories, recommended data sharing practices, etc.)

Platforms — Platform for Experimental Collaborative Ethnography

DARIAH
- Advances arts and humanities scholarship in Europe
- Offers repositories, tools and consulting

Project Bamboo
- Attempted to advance digital humanities in the US (eventually failed)
- Created the dirt repository

CLARIN
- Advances humanities research about language in Europe
- Offers repositories, tools, and consulting

ICPSR
- Advances social and behavioral science for the public good
- Offers tools and consulting

GESIS
- Advances statistical and quantitative social science
- Offers repositories, tools, and consulting

California Digital Library
- Advances research in the University of California system
- Offers repositories, tools, consulting (DMPTool)

The EHCSSRII would offer research infrastructure governance for vetting, networking, and sustaining frameworks, tools, and services to advance CHI.

EMPIRICAL HUMANITIES AND CREATIVE SOCIAL SCIENCE RESEARCH INFRASTRUCTURE INITIATIVE

Organizational Support
- Connect researchers with relevant organizations already addressing data concerns, such as the Research Data Alliance and DHCommons
- Advocate for the development of research infrastructure attuned to the special needs of EH & CSS
- Develop model documents, vetted by organizations like the Research Data Alliance, that outline the value of funding, designing, and sustaining research data infrastructure
- Devise new mechanisms for crediting scholars that participate in the design and implementation of digital research infrastructure
- Develop and/or vet certifications for following recommended data practices, which could be publicized to potential funders

Technical Services and Support
- Create a data net for networking repositories that hold empirical humanities and creative social sciences data
- Offer repository space for digital tools, services, and platforms built by digital humanists
- Offer access to services such as persistent identifiers, registering data types, data management software, and a SPARQL endpoint/other querying functionality
- Maintain a directory of tools and services that support humanists such as recommendations for data provenance, nano-publications, metadata, data licensing, and best practices for data citation
- Host workshops and offer training at relevant societal conferences

Critical Data Education
- Encourage critical reflection on data practices and advance thinking and research about how to apply hermeneutic sensibilities to data practices and data infrastructure design
- Develop model curriculums for critical data literacy
- Offer workshops on critical insights on data infrastructure at RDA meetings
- Prepare the next generations of digital humanists to sustain the EHCSSRII, contributing to organizational and technical support efforts