

Tales of the Field: Building Small Science Cyberinfrastructure

Andrea Wiggins, Syracuse University School of Information Studies

An interdisciplinary community of researchers has started to coalesce around the study of free/libre open source software (FLOSS) development. The research community is in many ways a reflection of the phenomenon of FLOSS practices in both social and technological respects. Researchers around the world are contributing to the development of cyberinfrastructure to support the virtuous cycle of academic research on FLOSS, and many share the open source community's values that support transparency and democratic participation. As community ties develop, new collaborations have spurred the development of shared resources and increased venues for interaction among researchers. As a result, several repositories provide ready access to curated, research-ready data, working paper repositories provide a means for disseminating early results, and a variety of analysis scripts and workflows connecting the data sets and literature are freely available. Despite these apparently favorable conditions for research collaboration, adoption of the tools and practices associated with this mode of research has been slow as yet.

This session will discuss the current state of small science cyberinfrastructure for FLOSS research, a vision for cyberinfrastructure-enabled research in this research community, and the perceived barriers to achieving this vision. The key issues observed to date seem to stem from the challenges of pre-paradigmatic small science research. While the community is willing and able to share assets, the interdisciplinary nature of the research prevents a concerted effort to define and pursue "grand challenges." In practical terms, this means that reuse of analysis tools, for example, is reduced because of the lack of convergence on key constructs and theoretical models. Researchers from software engineering, information systems, and even anthropology may examine the same construct, such as FLOSS project success, but will likely proceed from different epistemologies, utilize different data sources, identify different independent variables with varying operationalizations, and employ different research methodologies.

At the same time, while the barrier to entry to membership in this research community is significantly reduced by the increasing availability of research data and literature on FLOSS, the requisite technical skills to benefit from the growing FLOSS cyberinfrastructure are only slightly reduced for new members to the community. In addition, any movement toward standardization introduces a new source of disincentive to individual members who must abandon current localized research group practices if they are to adopt new research practices in the name of broader community benefit. FLOSS researchers have nonetheless built a significant foundation for cyberinfrastructure-enabled research, motivated by the potential greater benefit of shared resources.

While none of these confounding factors are new or unique to FLOSS research, the small size and interdisciplinary nature of the active research community brings greater salience to concerns over the mediating effects of the scale of participation. In the decentralized and phenomenology-driven FLOSS research community, creating and maintaining cyberinfrastructure is a substantial effort for a small number of participants. For small sciences, critical mass of participation may be the most significant factor in creating a viable cyberinfrastructure.