# Digital Scholarship Group: Inception and Startup

A Report on Startup, Funding, Staffing, and Projects, 2014–2017

October, 2017

# 0. About This Document

This document reports on the first phase of development (January 2014-June 2017) for the Digital Scholarship Group in the Northeastern University Library, including its initial launch, early planning and policy development, and first-stage tool and infrastructure development. It also describes the group's staffing and funding models, working partnerships, and major projects. The appendix lists grant proposals, awards, events, and teaching/consultation activity.

# 1. Rationale and Mission

DSG is an applied research group with an overall focus on exploring the increasing and distinctive role for libraries in digitally mediated processes of scholarly research and dissemination. We study and support processes at many phases of the research lifecycle, including data creation and curation, digital information representation and modeling, project planning and management, education and training, and technologies of scholarly publishing. Members of the group maintain expertise in, and actively contribute to, research in information standards, tool architectures, information modeling, data curation, linked open data, data analysis and visualization, as well as many more specific areas of relevance (such as digital scholarly editing, open access publishing, geospatial data, etc.). DSG's areas of expertise complement and engage with research directions arising from faculty-led initiatives such as the NULab, the NUVis Consortium, and the Humanities Center, as well as individual research projects by faculty and graduate students. DSG's work also complements initiatives within the colleges (especially CSSH, CAMD) to build graduate programs in digital humanities with an experiential dimension; in partnership with NULab, DSG provides a strong professional mentoring environment for graduate students, including substantive jobs, contribution to research efforts and grant-funded projects, and intellectual community.

# 2. Startup Process

DSG opened its doors in January 2014. The group's first phase of operation was spent in three main areas of effort: community development, policy development, and infrastructure development, each of which is described in more detail below.

## 2.1 Community Development

The process of community development at the outset involved creating a visible public presence for DSG, building working relationships with other relevant groups, and developing an initial set of programs of support. The specific activities during this period included:
- With the NULab, establishing Digital Humanities Open Office Hours: an informal weekly brown-bag event, held in the Digital Scholarship Commons, which combines informal presentations (from within and outside the NEU community) and discussion. Novices are encouraged to bring questions, and the venue also serves as a place to rehearse conference presentations, share interesting readings, and trouble-shoot interesting technical challenges.
- Establish a curriculum of workshops on topics in demand by the digital humanities community: TEI and XML-related technologies, Omeka, grant management, project management tools, data visualization, GIS. (A full list of topics and events is included in the appendix.)
- With the NULab, develop a program of visiting scholars and interns; we have hosted multiple interns from Northeastern and from the Simmons library school, and visiting scholars from Germany, Sweden, and Australia. This program is documented at http://dsg.neu.edu/about/jobs/.

- With the NULab, offer periodic open house/showcase events, including an annual "fall welcome" event.
- Develop the DSG web site as a space to present our resources to the Northeastern community and also as a space to document our activities and make them visible to peer organizations.

## 2.2 Policy Development

DSG's early policy development during our startup phase focused on three high-priority areas:

- **Project handling:** We developed an internal set of definitions for what we meant by "a DSG project" including a consideration of what types of projects and functionalities we could realistically support, what kinds of commitments we could make, what boundaries to set concerning types of content and the range of users to support. We also developed protocols for project intake (including what information to be requested at each stage of the process, planning meetings and who should attend, boilerplate communications), project tracking (including a tracking database and associated submission forms), and project closure (including decisions about what information to archive and how).
- **Documentation:** We developed protocols for documentation, including a shared understanding of what kinds of information should be documented, at what level of detail and for what audience, and where different kinds of documentation should be stored. We set up initial systems based on these discussions that make use of the DSG wiki (for internal policies and procedures, boilerplate communications, and draft training materials), Podio (for project records and actions, data management planning, data about grant applications and events), the DSG web site (for public-facing policies and training materials), Google Docs (for drafts of long documents such as specifications that require a variety of input and flexible co-authoring) and GitHub (for code and code-related documentation).
- **Data management planning:** Working with the library's Research Data Management Librarian and the Digital Repository Manager, we developed a data management planning instrument adapted from several existing data management planning tools, to use with DSG projects as part of the project startup process and also at intervals during the digital project lifecycle. Our initial intention was that all DSG projects should have a regularly updated data management plan on record, both to support DSG's data curation responsibilities and also to be prepared for potential funding applications. Over time, as the CERES tool set has evolved, we have revisited this requirement and focused our efforts on larger projects with more complex data, but we encourage all DSG projects to maintain a current data management plan. Our data management planning guidance is documented at http://dsg.neu.edu/services/development-management/data-management-plans/.

## 2.3 Infrastructure Development

The third major area of effort in the first year was the design and provision of infrastructure for digital scholarly projects and publications. As part of our initial planning, we established support for Open Journal Systems (OJS, an open-access journal publishing platform) and for Omeka (a content management system aimed at cultural heritage practitioners), including documentation and training materials and sandbox instances for experimentation. The majority of our focus, however, was on what would become the CERES publishing toolkit, initially called "DRS Plus" and then "DRS Project Toolkit" (for more details, see below and also http://dsg.neu.edu/ceres/). The overall goal of this service is to support the creation of complex scholarly digital publications and research environments, while storing the primary digital assets for these projects in the Northeastern University Digital Repository Service (DRS), and enabling the DSG to support a wide range of digital projects without having to provide individualized support for idiosyncratic tools and technical architectures. This toolkit was envisioned from the start as a long-term development effort that would begin with a simple set of features and

expand over time, driven by the needs of projects seeking to use it. We held a variety of design meetings during the course of the year to consider both the planned user functionality and the underlying architecture, and also planned out a pilot process to include a call for proposals, an outreach plan, and a development timeline. The initial feature set was implemented as a WordPress plugin that interacts with a repository API (built in Hydra/Samvera). It focused on drawing file thumbnails and metadata dynamically from the DRS, and on supporting the creation of exhibits and dynamic pages.

Following this planning process, at the start of our second year (2015) we began the pilot program for the toolkit. An initial CFP was issued in February 2015, which yielded 11 proposals from diverse areas: History, Archives, Political Science, Jewish Studies, Criminal Justice, DMC Studios, Engineering, Nursing, Professional Studies, Music. Of these we selected three for immediate implementation, and waitlisted three more; one further project was implemented as a straight DRS project without requiring a Toolkit site. We had also identified three other projects that were already a priority for the pilot, for a total of 9 projects selected. Of the projects not selected, several were too ambitious for the current state of the Toolkit but represented the start of a longer conversation. A few did not meet our criteria for DSG projects (either because they represented private data-gathering rather than scholarly research, or because they were a poor fit for the Toolkit overall. Development of the toolkit began in early 2015. We held project intake meetings in early summer and began training project teams to use the toolkit in September 2015. Five projects from this initial pilot set are now publicly visible; one is still in active development by their project teams; details are provided in the appendix to this document.

Following this pilot year, in early 2016 and again in early 2017 we issued the second and third annual calls for proposals. In 2016 we gave particular emphasis to projects that would allow us to design and test new mapping and timeline features identified in the pilot phase as high priorities for inclusion; in 2017 we focused on pedagogical projects that would use CERES in the classroom. With the basic technical infrastructure in place, we were able to offer CERES as a "client" project in a graduate CCIS course in which a student team developed code modules to support these new features, with oversight from our development staff. We also expanded our training program for CERES project teams. To support this training we developed a fuller set of documentation and training materials which have now been given their own CERES site (http://cerestoolkit.dsg.northeastern.edu).

# 3. Staffing and Funding

DSG is designed as an applied research group with a variety of different types of technical and research expertise. Working in partnership with faculty and project teams, members of DSG need to be able to reflect and mediate knowledge about technical systems, information standards, digital tools, and digital scholarly methods. Because this work entails such a wide variety of expertise domains, members of the group also need to be adept at learning from one another and dovetailing their own knowledge with that of the group. The group is designed to include staff with backgrounds in library and information science, digital humanities scholarship, and programming and technical development.

DSG's initial staffing was drawn in part from existing library positions with a strong digital emphasis and in part from the Women Writers Project, whose inclusion as a foundational project at DSG's launch brought with it several federal grants and two staff positions. In addition to its own core staff, DSG also includes dotted-line relationships that bring other relevant positions into the conversation, including the university archivist, the research data management librarian, the scholarly communications librarian, and members of Library Technology Services. During the initial development period described here, the initial staff group was augmented with additional positions including a GIS specialist, a data visualization specialist, a web applications programmer, and a semantic data specialist; some of these were explicitly short-term positions (funded by external grants) and others were longer-term positions.

Finally, in addition to DSG's staff positions, the group includes several student positions which provide crucial support for training, documentation, project management, and contributions of effort to specific DSG projects.

Funding for DSG staff during this first phase has come from a variety of sources. Approximately a third of the positions are funded as permanent library staff lines. Another third are funded from a combination of grant funding and project-related funding, including income from Women Writers Online. The remaining staff are funded by the university as short-term pilot positions, with the goal of moving these over time into permanent library staff lines. Students are funded from the library budget, supplemented by income from workshops and consulting. Grant funding plays an important role in DSG's finances. During this initial period it has enabled us to support essential staff positions for which firm institutional funding does not yet exist, and also student positions. In the longer term, grant funding will be used more for short-term specialized roles required by specific initiatives.

## 3.1 Staff

The following specific areas of expertise have been given the highest priority in the initial staffing of DSG:
- Information standards and metadata standards
- Data modeling
- XML technologies and tools
- Geospatial approaches
- Data visualization
- Management and sustainability of large-scale, long-term digital projects
- Linked open data and semantic web technologies
- Repository architectures
- Digital project management
- Data curation

The specific staff and student positions within DSG, as of the completion of this first development phase, are as follows:
- Director (Julia Flanders)
- Assistant Director (Amanda Rust)
- XML Applications Developer (Ashley Clark)
- XML Programmer-Analyst (Syd Bauman)
- GIS Specialist (Bahare Sanaie-Movahed)
- Data Analytics and Visualization Specialist (Steven Braun)
- Assistant Director, Women Writers Project (Sarah Connell)
- Semantic Data Specialist (Jeffrey Bourns)
- DSG Coordinator (student position)
- Charon Coordinator (student position)
- CERES Classroom Coordinator (student position)

Neighboring staff positions with whom DSG works closely:
- Digital Repository Manager (Sarah Sweeney)
- Research Data Management Librarian (Jennifer Ferguson)
- Director, Scholarly Communication & Digital Publishing (Hillary Corbett)
- Head, Archives and Special Collections (Giordana Mecagni)
- Web Developer and Designer, Library Technology Services (open position)
- Associate Dean for Digital Strategies and Services (Patrick Yott)
- Systems Coordinator, Library Technology Services (Karl Yee)
- Senior Library Developer, Library Technology Services (David Cliff)

## 3.2 Partnerships Within the Library

DSG has important working partnerships with several groups in the library:

**Research and Instruction:** DSG works with subject liaisons, who are often the first point of contact with faculty concerning their research, and are a direct resource for departments. Members of R&I attend DSG meetings on CERES projects (including our overall program planning and specific project meetings). DSG also works with R&I to plan workshops for library staff (e.g. in GIS and data visualization). R&I also collaborates with DSG on specific initiatives, including:

- Collaboration with Art & Design librarian on data visualization and CAMD DRS use
- Planning of Wikipedia classes in partnership with Communications Librarian
- Hosting a Wikipedia Visiting Scholar in Women's and Gender Studies, involving both Communications and WGSS librarians

**DMC Studios:** DSG works with the DMC studios to design digitization work flows for joint projects, and to draw on their expertise

**The Research Services Group:** DSG works closely with the Research Services Group on data and project curation policy development and support for faculty grant planning. Specific projects include joint work on the Ocean Genome Legacy (OGL) data upload project involving collaboration between DRS Manager and Biology and Research Data Management Librarians. The DSG also works with each of Research Services Group's three members to provide access to research services and incorporate research lifecycle principles and practices into regular DSG activity:

- **Repository Services:** The DRS is an essential part of the DSG services infrastructure, providing repository support and storage for the more than twenty CERES sites. The repository manager also provides metadata guidance for project teams, assists in depositing project materials, and helps to organize and manage CERES projects.
- **Scholarly Communication**: DSG works closely with the Director of Scholarly Communication and Digital Publishing on strategic planning concerning platforms for digital scholarly publishing, and to provide support for Open Journal Systems and for the Scholarly Communication & Digital Publishing Program
- **Data Management**: The DSG works closely with Jen Ferguson to develop and implement data management planning strategies for digital scholarly projects, and to provide guidance and support for faculty grant-writing.

**Archives and Special Collections:** This group is a major partner in DRS and CERES development. The CERES Toolkit is used extensively by Archives staff and interns to create exhibits and collection portals, and many DRS upload features were developed with Archives input or by Archives request. DSG has also worked with Archives on grant proposal development.

**RADS:** DSG draws on RADS expertise in several important areas including billing support, licenses, and contracts (including the exploration of expanded data mining access to licensed digital collections), digital material cataloguing, and metadata expertise from MROG. We also look forward to working with the next Digital Production Coordinator to create and review digital project materials and produce guidelines for project teams engaged in digitization activity, and support.

## 3.3 Partnerships Outside the Library

DSG's closest partnership outside the library is with the NULab for Texts, Maps, and Networks. The two groups collaborate on events, workshops, and visiting scholar programs, and NULab also serves as an important focal point for faculty research in digital humanities and computational social sciences and hence a point of contact for DSG's support. DSG staff provide mentoring and expertise to graduate

students in the NULab's Digital Humanities Certificate program, and DSG trains and oversees graduate and undergraduate student workers on NULab faculty research projects. DSG is also a partner in the NUVis Consortium (http://nuvis.northeastern.edu) of which Steven Braun is the Associate Director, working with co-directors from CCIS and CAMD; NUVis sponsors talks, symposia, and other community-building events on interdisciplinary data visualization.

DSG also has important collaborative partnerships with the digital humanities and digital library community outside of Northeastern, in several areas:

- Design and development of Samvera/Fedora-based infrastructure for complex digital publications and research projects: DSG is in regular communication with research groups including the MLA CORE, the Canadian Writers Research Collaboratory, Washington University St. Louis, and the recent NHPRC/Mellon initiative to support federated infrastructure for digital scholarly editions
- Hosting of international digital humanities projects and publications: DSG hosts the TEI Archiving, Publishing, and Access Service (TAPAS), and hosts the editorial and production work for the Digital Humanities Quarterly journal (including an OJS installation for workflow tracking; the actual publication is hosted on ADHO servers).

DSG staff also contribute to international standards work and to the work of professional organizations: DSG contributes effort to the TEI Consortium Technical Council, the Digital Commonwealth Repository Committee, New England Wikimedians, and ACRL.

# 4. Major Projects

DSG's programs and services during this startup period were focused around several major projects, some of which were catalysts for external funding, and all of which provided ways to frame and scope our activities.

## 4.1 CERES and Charon Projects

As described above, CERES (the Community-Enhanced Repository for Engaged Scholarship) is the repository-based infrastructure developed by DSG and LTS through which our support for digital projects is largely focused. CERES currently consists of a Hydra/Samvera "head" that exposes data from the DRS for interaction and publication via a customized WordPress front end. The WordPress interface supports the creation of exhibits, narratives, and exploratory interfaces that draw metadata and digital objects dynamically from the DRS at the item or collection level. This feature set has expanded over time to include dynamically generated maps and timelines. DSG also created extensive documentation and training programs to support CERES project teams, and also to support the use of CERES in the classroom. In the future, DSG will continue to add features and may also add other publication tools as options within CERES. CERES currently houses over 25 projects at various stages of completion; 6 are completed and publicly visible on the DSG site (http://dsg.neu.edu/projects/#toolkit). A full list is included in the appendix.

Building on CERES, the next phase of development has been the planning of Charon, a platform to support a complete editorial workflow for complex document-based projects, including transcription, XML encoding, translation, proofreading and editorial review, and annotation. Charon will be implemented in Hydra/Samvera, and will assume TEI data in the first instance. A full specification was circulated in spring/summer 2017 to members of the Hydra/Samvera community and to research partners at peer organizations. Development is being undertaken in AY17-18. We have identified three

pilot projects focused on different types of complex TEI data, which will be used to test Charon during its development:

- The Jesuit Relations project: a digital documentary edition of a 40-volume 17th-century record of Jesuit missions to the New World.
- The American Indigenous Language Education and Sustainability project (AILES): a digital collection of Cherokee-language documents to serve as the basis for language learning and cultural study.
- The Early Caribbean Digital Archive (ECDA): a large-scale digital archive of published and unpublished documents from the 17th-19th century Caribbean region, with the goal of supporting annotation and discussion by a wide collaborative scholarly community of researchers that also includes community members and non-academics.

## 4.2 Funded Research and Development

DSG seeks external funding for special projects; funding pays for student time and for staff who are currently on soft money. We also collaborate with faculty members on grant proposals to support (or hasten) the development of DSG's research and publishing infrastructure. Since DSG's founding in January 2014 we have submitted 14 external grant proposals, of which four have been awarded; two others were given preliminary approval at the first round but we decided not to submit a full proposal. An additional proposal had been submitted prior to DSG's founding and was awarded in early 2014. We have also been partners on two successful Tier One internal awards. The awarded grants are:

**IMLS National Forum on Design for Diversity (DSG; $99,822)**: a two-year award funding two open forums and a teaching and learning toolkit on approaches to the design of information systems for libraries and cultural heritage institutions in ways that reflect goals of diversity and social justice. The toolkit will include case studies, course modules, an annotated bibliography, and a listing of other resources. More at http://dsg.neu.edu/research/design-for-diversity/.

**NEH Collaborative Research: Intertextual Networks (WWP; $290,000):** a three-year award funding the identification of all bibliographic sources referenced from Women Writers Project texts, and the linking of in-text citations, quotations, and other references to those sources; this grant also supports an extensive set of individual research projects on intertextuality, which will be published by the WWP in Women Writers in Context. More at http://www.wwp.northeastern.edu/research/projects/intertextuality/index.html.

**NEH Research and Development: TAPAS: Building an XML-aware Repository (TAPAS; $290,000):** a three-year award (extended) funding the development of an XML-aware repository back end to the TAPAS service.

**NEH Digital Humanities Advancement: TAPAS Classroom (TAPAS; $74,630):** an 18-month award funding the development of specialized stylesheets, validation tools, and training materials to support teaching TEI with TAPAS.

**NEH Digital Humanities Advancement: Mining Citation (DHQ; $74,123):** an 18-month award funding the completion of a centralized bibliography of digital humanities to support the DHQ journal and facilitate the study of digital humanities research and citation practices.

**Northeastern Tier One: Big Data and Text Curation (in partnership with the English Department and Computer Science; $50,000):** a 15-month award funding the exploration and development of natural language processing models for automated detection of genres and named entities in digital archives.

**Tier One: Word2Vec (in partnership with the English Department and Computer Science; $50,000):** a 15-month award funding the development of a prototype graphical user interface to support the use of word-embedding models of text analysis by novice users.

## 4.3 Outreach, Consultation, and Mentorship

A significant portion of DSG's effort is focused on supporting individual students and faculty, and on strengthening the Northeastern research and learning community in areas of digital scholarship, digital humanities, and computational social sciences. Specific programs and activities include:

**GIS and data visualization consultation:** Bahare Sanaie-Movahed and Steven Braun provide consultation to faculty and students, including regular weekly office hours, and are available to take on larger-scale projects supported by grant funding. More at http://subjectguides.lib.neu.edu/gis-dataviz.

**Public workshops:** DSG offers regular workshops on GIS, data visualization, TEI and XML technologies, using the DRS, data management, Wikipedia editing, and grant-writing. Most are open to the public (aside from purely internal topics such as using the DRS).

**DH Open Office Hours:** With the NULab, DSG co-hosts this weekly brown-bag session, aimed at fielding impromptu questions, hosting informal presentations by visitors and members of the NEU community, discussing readings, and trouble-shooting interesting problems in digital humanities.

**Research assistantships, student positions, internships, and visiting scholars:** DSG oversees a varied and substantive range of opportunities for students, both paid and unpaid. DSG and its grant-funded projects offer several research assistant positions and hourly student positions (aimed at graduate students. Students are embedded as colleagues in established digital humanities projects and participate in writing documentation, grant-proposals, research reports, and publications, as well as contributing to technical work including TEI/XML encoding, metadata creation, programming, data visualization, and georeferencing. DSG also offers internships and visiting scholar programs that attract graduate and undergraduate students from Northeastern, Simmons, and other area institutions. Specific internship programs include Public History, GIS, and projects arising from the DRS and CERES. More at http://dsg.neu.edu/about/jobs/.

**Data sets for teaching and research:** DSG provides data from its projects to faculty and researchers at Northeastern and other institutions, including the NULab's annual "Speed Data-ing" event, an information visualization MOOC at Indiana University, a newly hired Sociology professor seeking data for a class assignment, and two Tier One grants. DSG project data is often valuable as teaching or training data because of its level of curation and markup.

# Appendix

# Active DSG Projects as of 2017

## CERES Projects

Projects using our home-grown WordPress theme and DRS API, most projects gathered through a public yearly Call for Proposals.

## Live Projects

**Boston's Latino/a Community History**
http://latinohistory.library.northeastern.edu/
Boston's Latino/a Community History is a collection of more than 40,000 digitized photographs and documents from Northeastern University's Archives and Special Collections' Inquilinos Boricuas en Acción and La Alianza Hispana collections.

**The Catskills Institute**
http://catskillsinstitute.northeastern.edu
The Catskills Institute was created by the organizing committee of the History of the Catskills conference that was held in Woodridge, New York on Labor Day Weekend 1995..

**DMC Studios Showcase**
http://dmcstudios.library.northeastern.edu/
The DMC Studios Showcase features student audio and video recordings and 3D printing projects created in the Library's state-of-the-art facilities.

**Latin American Jewish Writers and Artists**
http://latinjewisharts.northeastern.edu
Stephen Sadow's interviews with Latin American Writers and Artists is a series of 12 video interviews recorded in Argentina by Professor Sadow himself in 2012, along with several works written about or by the artists.

**"Picturing the World" Gallery**
http://arader.library.northeastern.edu
The Library's "Picturing the World" gallery is a collection of maps, watercolors, and Audubon prints.

**A Proud Past: A History of Boston-Bouvé College, 1913-1981**
http://aproudpast.library.northeastern.edu
A Proud Past is a history of the Boston-Bouvé College as told through photographs from Northeastern University's Archives and Special Collections.

## Projects In Development

- African American Institute Archive
- Dragon Prayer Book (Erika Boeckeler, Department of English)
- Civil Rights and Restorative Justice Project (NULawLab, School of Law)
- Early Black Boston Digital Almanac (Nicole Aljoe, Department of English)
- Early Caribbean Digital Archive (upcoming transition to Toolkit site)
- Henry David Thoreau Journal Drawings (Kathleen Kelley, Department of English)
- Holocaust Awareness Week Programming (Holocaust Awareness Committee/Jewish Studies)
- Spectrum Literary Magazine Archive

# DSG-Supported Projects

DSG-supported projects are projects that belong directly to the Digital Scholarship Group, and often have their own staff within the group.

**Digital Humanities Quarterly (DHQ)**

Digital Humanities Quarterly (DHQ) is an open-access, peer-reviewed, digital journal covering all aspects of digital media in the humanities.

**Digital Feminist Commons**
The Digital Feminist Commons, which is funded by the Northeastern Humanities Center and the Association for Computers and the Humanities, is a space to generate cross-disciplinary discussions on the intersection of feminist and digital research.

**Women Writers Project (WWP)**
The Women Writers Project is a long-term research project devoted to early modern women's writing and electronic text encoding.

**TEI Archiving Publishing and Access Service (TAPAS)**
TAPAS is the TEI Archiving Publishing and Access Service for scholars and other creators of TEI data who need a place to publish their materials in different forms and ensure it remains accessible over time.

## DSG/NULab Projects

The DSG works closely with the NULab to offer tools and support for faculty projects. The following projects are built and staffed by members of the NULab with support for grant-writing, software, and development work from the DSG.

**Early Caribbean Digital Archive (ECDA)**
The Early Caribbean Digital Archive (ECDA) is a highly interactive digital scholarly lab for the collaborative research and study of pre-20th century Caribbean literature. The ECDA seeks to engage scholars and students in a shared, critical study of the textual, material, and cultural histories of the Caribbean by providing them innovative digital technologies and platforms for generating new and understudied knowledges of the Caribbean's rich body of materials. The ECDA is in the process of migrating into CERES.

**Viral Texts**
In its first phase (2012-2014), Viral Texts focused on developing its text reuse discovery algorithms and investigating reprinting in the nineteenth-century United States. Links to publications from this phase can be found on our publications page. In 2015-2016, the Viral Texts team will focus on improving the project's algorithms and investigating international reprinting among English-language newspapers in the United States, United Kingdom, and Australia.

**Our Marathon**
Our Marathon is a crowd-sourced archive of pictures, videos, stories, and even social media related to the Boston Marathon; the bombing on April 15, 2013. Our Marathon is in the process of migrating into CERES.

# Project Statistics

## Departmental Involvement

### Requesting Department for CERES Installations
**Total**                                        **25**

| Archives & Special Collections (Snell Library) | 9 |
|---|---|
| College of Social Science & Humanities (CSSH) | 4 |
| English | 4 |
| Scholarly Communication & Publishing (Snell Library) | 2 |
| Office of the Dean (Snell Library) | 1 |
| Sociology & Anthropology | 1 |
| NuLawLab | 1 |
| Bouve College of Health Sciences | 1 |
| College of Arts, Media, & Design (CAMD) | 1 |
| NULab | 1 |
| Studios (Snell Library) | 1 |
| Language, Literatures, and Cultures | 1 |
| Jewish Studies | 1 |

## Department Sponsoring Projects

| **Total** | **27** |
|---|---|
| College of Social Science & Humanities (CSSH) | 15 |
| English | 11 |
| Digital Scholarship Group (Snell Library) | 4 |
| NULab | 2 |
| Jewish Studies | 2 |
| Sociology & Anthropology | 1 |
| NuLawLab | 1 |
| African-American Institute (AAI) | 1 |
| School of Law | 1 |
| Humanities Center | 1 |
| Archives & Special Collections (Snell Library) | 1 |
| Studios (Snell Library) | 1 |
| Language, Literatures, and Cultures | 1 |
| Political Science | 1 |

## Stages of Development

| **Total** | **27 projects** |
|---|---|
| Active Development | 19 projects |
| Proposed | 5 projects |

| Dormant | 2 projects |
| Ongoing Maintenance | 1 projects |

# Grant Information

Grants applied for but not received:

- NEH Digital Implementation (February 2014): Implement Bookworm for DSG projects
- NEH Collaborative Research (April 2014): WWP Intertextual Networks
- NEH Humanities Collections and Reference Resources (July 2014): ECDA
- NEH Collaborative Research (December 2014): WWP Genre and Textual Materiality
- NEH Research and Development (June 2015): Repository development
- NEH Digital Implementation (February 2016): NULawMaps
- Software and Data Carpentry Workshops (April 2016): UMass Worcester NN/LM, New England Region
- NEH Digital Advancement (January 2017): Visualizing TAPAS

In addition, three IMLS preliminary proposals were submitted, of which two were approved for a full application; we submitted one full application, which was funded (see below).

Grants awarded:

- NEH Research and Development ($290,000, 2014-18): TAPAS, Building an XML-Aware Repository
- IMLS National Forum ($99,922), 2017-18): DSG, Design for Diversity
- NEH Collaborative Research ($289,000, 2016-19): WWP, Intertextual Networks
- NEH Startup ($74,630, 2015-17): TAPAS, TAPAS Classroom
- NEH Startup ($74,123, 2017-19): DHQ, Mining Digital Humanities Citation
- Northeastern Tier One ($50,000, 2015-16): ECDA/WWP, Big Data and Text Curation
- Northeastern Tier One ($50,000, 2017-18): WWP/CCIS, Word2Vec

Applications/awards per agency:

| Total | 15 applications / 7 awards |
| NEH | 11 applications /4 awards |
| IMLS | 1 application /1 award |
| NLM | 1 application /0 awards |
| Northeastern | 2 applications /2 awards |

Amount awarded per agency:

| Total | $827,575 |
| NEH | $727,753 |
| IMLS | $99,822 |
| Northeastern | $100,000 |

# Other DSG Activity

[Totals are not a direct sum of category-level subtotals, since items may appear in more than one category.]

# Public Events

## Public Events by Type

| | |
|---|---|
| **Total** | **158** |
| Discussion / Reading Group | 69 |
| Informational / Demo Presentation | 56 |
| One-on-One Drop-in (no formal agenda) | 12 |
| Extended Workshop (1 or more days) | 11 |
| Working Group | 7 |
| Internal (Library-only) | 5 |
| DissCo | 3 |
| Hands-on Practicum Workshop | 2 |

## Attendance by month

| | |
|---|---|
| **Total** | **1,409** |
| February 2017 | 46 |
| January 2017 | 42 |
| December 2016 | 8 |
| November 2016 | 162 |
| October 2016 | 194 |
| September 2016 | 30 |
| April 2016 | 15 |
| March 2016 | 57 |
| February 2016 | 53 |
| January 2016 | 32 |
| December 2015 | 33 |

## Attendance by Event Type

| | |
|---|---|
| **Total** | **1,409** |
| Discussion / Reading Group | 672 |
| Informational / Demo Presentation | 571 |
| Working Group | 66 |
| Extended Workshop (1 or more days) | 49 |
| DissCo | 33 |
| One-on-One Drop-in (no formal agenda) | 31 |
| Internal (Library-only) | 31 |

| Hands-on Practicum Workshop | 27 |
|---|---|

## Teaching and Consultations

### Consultations vs Classroom Work

| Total | 373 |
|---|---|
| Consultation | 310 |
| Custom Workshop, Class Visit, or Demo | 63 |

### Audience (all consults & workshops)

| Total | 360 |
|---|---|
| Graduate | 184 |
| Faculty/staff | 104 |
| Undergraduate | 97 |

### Special Topics for Consultations

| Total | 310 |
|---|---|
| GIS | 204 |
| Data Visualization | 54 |
| DRS | 33 |
| Grant (planning, writing, submitting) | 5 |

### Consultation Attendance by Unit

| Total | 466 |
|---|---|
| College of Social Science & Humanities (CSSH) | 164 |
| College of Science (COS) | 50 |
| College of Professional Studies (CPS) | 38 |
| College of Arts, Media, & Design (CAMD) | 37 |
| College of Engineering (COE) | 27 |
| English | 27 |
| Bouve College of Health Sciences | 25 |
| Center for Resilience Studies | 23 |
| College of Computer & Information Science (CCIS) | 22 |
| Political Science | 18 |
| D'Amore-McKim School of Business (DMSB) | 17 |

| Economics | 17 |
| Marine and Environmental Sciences | 16 |
| Behavioral Neuroscience | 15 |
| School of Public Policy and Urban Affairs | 14 |

## Special Topics for Classroom Work

| **Total** | **63** |
|---|---|
| GIS | 19 |
| DRS | 19 |
| Data Visualization | 9 |

## Classroom Work Attendance by Unit

| **Total** | **757** |
|---|---|
| College of Social Science & Humanities (CSSH) | 282 |
| English | 251 |
| College of Arts, Media, & Design (CAMD) | 155 |
| School of Journalism | 93 |
| D'Amore-McKim School of Business (DMSB) | 71 |
| College of Computer & Information Science (CCIS) | 60 |
| Architecture | 37 |
| College of Science (COS) | 32 |
| Communications | 31 |
| Music | 30 |
| College of Engineering (COE) | 30 |
| Studios (Snell Library) | 24 |
| Political Science | 18 |
| Archives & Special Collections (Snell Library) | 17 |
| Marketing and Communications (Northeastern) | 13 |