Integrating Humanities and Science: the Scriptospatial Visualization Interface

Fenella G. France
frfr@loc.gov
Preservation Research and Testing Division
Library of Congress, United States of America

Alberto Campagnolo
acamp@loc.gov
Preservation Research and Testing Division
Library of Congress, United States of America

Introduction

Advances in technology and digital access have paved the way for the improved utilization and interpretation of scientific analyses of humanities materials for digital humanities studies. Integrating scientific analyses with humanities and curatorial knowledge (STEM: science, technology, engineering, and math, to STEAM: science, technology, engineering, **art**, and math) is a critical multidisciplinary approach for expanding the full potential of scientific techniques and technological advances, and realigning complementary disciplines that have been artificially segregated. Scientists and curators have exposed hidden and previously unknown contextual information within original source materials, such as changing "subjects to citizens" on the Rough Draft of the Declaration of Independence (Library of Congress, 2010). Hyperspectral imaging provides additional data layers by capturing images of documents in distinct narrow waveband regions of the visible and non-visible spectrum—from ultraviolet through visible to infrared. The cube of captured digital image files contains a wealth of information, but requires significant interpretation to process and analyze the collected data. Scientists, scholars, and students in both art and science disciplines have been collaborating to glean new information from historical manuscripts. The Library spectral imaging program includes a spectral reference database and integration of data from other non-invasive analytical techniques to create a full analytical mapping of heritage documents and objects for non-destructive analyses of collection materials (France, 2016)

Digital imaging capabilities allow researchers to characterize pigments and inks on the document, retrieve hidden and lost text, and illuminate production and creation methods. The range of data captured allows greater access to the information available from fragile historic documents, including the 1507 Waldseemüller World Map and the 1513 Ptolemy *Geographia*, where investigations revealed links to the same original printing location (France, 2016). *Scriptospatial* (a term originally coined by Toth and France to refer to the viewing of associated imaging and materiality data linked on an image of a historic document) refers to applying a spatial information system approach to documents, creating an interactive interface for scholars and scientists to interact with the object and the data. Scriptospatial representations of digital data from documents utilize an accurate coordinate system that links scientific and scholarly analyses to the creation of a new digital cultural object (DCO), allowing inferences to be drawn to generate new knowledge. This approach to viewing digital cultural materials in multiple layers applies an archaeological approach toward uncovering and interconnecting information strata of historic and modern documents. Scriptospatial mapping of documents with an accurate coordinate system allows the layering of scientific and scholarly analyses to the DCO. This allows inferences to be drawn to generate new knowledge through analysis of the data linked to spatial points (or areas). This approach to viewing the DCO applies a GIS methodology toward uncovering and interconnecting information layers of cultural heritage artefacts, just as in the case of archaeological strata. Utilizing an object-oriented approach in conjunction with the spatial data layers allows the mapping of spatial and temporal data with increasing complexity. Examining and explaining the physical, spectral and chemical properties of these historic materials permit scientists and scholars to link these scientific analyses to other data about the creation of the object.

Digital spectral imaging of cultural heritage objects at the Library of Congress has capitalized on over a decade of research and development into not only spectral imaging and processing, but also the development of standard spectral image products (France et al, 2010). Advances over the past decade by an experienced team have led to an advanced capability to study cultural heritage objects with a robust spectral imaging system that provides large-format, high quality images and standardized data output using advanced commercial off-the-shelf components (Christens-Barry et al, 2009).

Developing an object-oriented approach to data access and sharing requires integration of spectral imaging with data from other sources in a variety of formats. This requires effective spatial metadata to allow linkages to specific locations within the images. This is necessary not only to register locations on the same section of a manuscript leaf in various spectral bands, but also to link other images and transcriptions with the spectral images. Based on geospatial mapping and layering of data used to identify points on satellite images, the same technologies, work processes and skills can be applied to spectral images of manuscripts: A camera collecting images over a manuscript is similar to a satellite collecting geospatial data over the Earth. Using technologies developed for "geospatial" systems to link each point on the globe with images from earth resource satellites and data collected from other instruments, spectral imaging can link the "scriptospatial data" from each point on a manuscript with images from various imaging and scientific devices. This provides a standardized method to support links between images and data from the same object location.

With multiple data entries for samples, precisely defining the specific point where the sample or scientific data collection takes place is critical in comparing data from different research types or objects. For samples taken from a larger, non-uniform, heterogeneous object such as a manuscript, textile or painting, the spatial location of the sample point on the object must be defined to be able to integrate the data from various research tools. Spatial metadata elements will allow linkages to specific locations on an object, potentially within images of the objects. Scriptospatial data can serve as an interface for scientific dialogue in "one shared layer," linking data from various sources for in-depth studies and analyses of a specific research topic or object.

In many current research databases, the metadata elements for spatial location are not provided to capture detailed data on where an instrument collects data, or a sample is taken. This is not an issue for uniform, homogeneous samples of paints, pigments, media or other samples, but is critical for samples taken from a heterogeneous object like a painting, manuscript or textile. By defining a Cartesian coordinate system on an object or image of an object, as well as the degree of precision required, specific sample points on an object can be defined. This allows integration with other images of the same object and scientific samples from the same point.

In its Scriptospatial Visualization Initiative, the Library of Congress PRTD has capitalized on developments with geospatial systems to apply Thermopylae "i-spatial" support and Google Map tiling and data formatting to the integration of large and complex visual scriptospatial datasets populated with scientific data from various instruments, research topics or objects. This provides data access in "one shared layer" of scientific data. This is an important first step in capitalizing on the three decades of technology development by the GIS community to advance preservation science and cultural heritage data sharing and research. The additional unique component will be the layering of scholarly research interpretations and publications, enabling ease of access to a rich resource of data directly linked to the original object. This will reduce the challenges faced with searching through data that is not well catalogued, or yet searchable without this expanded document interpretation and linking of scholarly knowledge.

Integrated access to associated source material scientific data adds contextual value, provenance information, and can reveal non-visible information hidden within the original document. The Scriptospatial approach allows layering of scientific and scholarly or curatorial research data within one location, enhancing the analysis and depth of knowledge off the original document, and creating a more effective interface for interaction with historic materials (France et al, 2010). The innate layering of multiple sources of information of the Shared Canvas Model (Sanderson and Albritton, eds, 2013) onto one view of an object, coupled with the strong annotation capabilities of the Web Annotation Data Model (Sanderson et al, eds, 2016) raises interest in the possible integration of the Scriptospatial model with IIIF efforts (International Image Operability Framework). One of the challenges we face with the current data deluge, is how to effectively access, select and link appropriate data and information. Scriptospatial is a value-added data approach, creating cohesive structured management of multidisciplinary data. The authors have been engaging with other US, European and United Kingdom colleagues to create a cohesive integrated and collaborative approach to this visualization.

Bibliography

Christens-Barry, W. A., Boydston, K., France, F. G., Knox, K. T., Easton, R. L. Jr., and Toth, M. B. (2009). "Camera System for Multispectral Imaging of Documents." In

SPIE Proceedings Vol. 7249: Sensors, Cameras, and Systems for Industrial/Scientific Applications X, edited by Erik Bodegom and Valérie Nguyen, 7249:8–10. San Jose (CA). doi:10.1117/12.815374.

France, F. G. (2016). "Spectral Imaging: Capturing and Retrieving Information You Didn't Know Your Library Collections Contained." In *What Do We Lose When We Lose a Library? Proceedings of the Conference Held at the KU Leuven* 9-11 September 2015, edited by Lieve Watteeuw and Mel Collier, 189–97. Leuven: KU Leuven University Library.

France, F.G., Emery, D., and Toth, M. B. (2010). "The Convergence of Information Technology, Data, and Management in a Library Imaging Program." *Library Quarterly.* Special Edition: Digital Convergence: Libraries, Archives, and Museums in the Information Age 80 (1): 33–59.

International Image Interoperability Framework. (2016) "IIIF | International Image Interoperability Framework." http://iiif.io/ (accessed 01-11-2016).

Library of Congress. (2010). "Hyperspectral Imaging by Library of Congress Reveals Change Made by Thomas Jefferson in Original Declaration of Independence Draft." *Press Release*, July 2. https://www.loc.gov/today/pr/2010/10-161.html (accessed 01-11-2016).

Sanderson, R., and Albritton, B., eds. (2013). "Shared Canvas Data Model 1.0." http://iiif.io/model/shared-canvas/1.0/ (accessed 01-11-2016).

Sanderson, R., Ciccarese, P., and Young, B, eds. (2016). "Web Annotation Data Model." https://www.w3.org/TR/annotation-model/ (accessed 01-11-2016).