Final Performance Report

Report ID: 94343
Application Number: PW5036909
Project Director: David Lester (dlester@umd.edu)
Institution: University of Maryland, College Park
Reporting Period: 5/1/2010 - 7/31/2011
Report Due: 10/31/2011
Date Submitted: 10/31/2011

The Text-Image Linking Environment (TILE) was awarded in 2009 by an NEH Preservation and Access grant to create the next generation of technical infrastructure supporting image-based editions and electronic archives of humanities content. With diverse partners including the Maryland Institute for Technology in the Humanities (led by Doug Reside and Dave Lester), Indiana University Bloomington, the Royal Irish Academy, the University of Oregon, and Harvard's Center for Hellenic Studies. Despite the proliferation of image-based editions and archives, the linking of images and textual information remains a slow and frustrating process for editors and curators. TILE, proposed as an extension of the NEH-funded AXE image tagger (http://mith.umd.edu/AXE/), was intended to increase the ease and efficiency of linking images and text within a web interface.

Project Activities
In our initial proposal, TILE was proposed to have produced software interoperable with other popular tools and capable of producing TEI-compliant XML for linking image to text and image to image with some level of automation. TILE would put the image linking features of the newest version of the Text Encoding Standard (TEI P5) through it's first rigorous, "real world" test, and, at the close of the project, provide to the TEI community a list of suggestions for improving the standard to make it more robust and effective. We intended to extend the functionality of AXE to:

- Semi-automated creation of links between transcriptions and images of the materials from which the transcriptions were made. Using a form of optical character recognition, TILE would recognize lines in a page image and link them to a pre-existing textual transcription.
- Annotation of any area of the image selected by the user with a controlled vocabulary
- Application of editorial annotations to any area of an image
- Supporting linking for non-horizontal, non-rectangular areas of source images
- Creation of links between different, non-contiguous areas of primary source images

Between May 1, 2009 and January 15, 2010, Dr. Doug Reside supervised all project related work through a 10% salary allocation. This included the assignment of Grant Dickie as web programmer, and Elizabeth Kvernan who acted as graphic designer and consultant in the creation of the software design plan and implementation of the development of the TILE software. On January 15, 2010, Dr. Reside left the University of Maryland and Dave Lester, Assistant Director of MITH, who had been brought into the project in October of 2009, began to supervise the project work. Mr. Lester was charged with completing all tasks as projected including finalizing project work, and completing this final performance report.

Most of the adjustments to TILE's project timeline were a result of staffing changes, and did not impact the final products of the grant. In order to complete the required work and use allocated funds, a no-cost extension was requested and granted for the TILE grant, which concluded July 31st, 2011.

Changes in project personnel at MITH, including the transition of Dave Lester as MITH Project Director in place of Doug Reside required up-front time and effort. During the same period, MITH was reconfiguring and growing its own staff, including the addition of another software developer and software architect. This additional staff, although often not working directly on TILE, established a local technical community at MITH to formalize our own software development process, including the establishment of the Lighthouse bug tracker and a more agile model of development. This extra capacity proved to be important during the final months

of the grant period and helped complete the development of the software. Other personnel changes included the movement of Dot Porter from the University of Kentucky, to Trinity, and then to Indiana University. Significantly, due to their own internal staffing changes, the roles of the University of Oregon and Harvard University declined through the lifecycle of the grant.

Project Accomplishments:
TILE has met some goals as defined in the initial proposal. These goals with the outlined accomplishments are listed below in table format for ease of reading.

Year One Goals:	Actual Accomplishments:
Initial Project Meeting	Cropping and re-arranging areas of an image. Linking portions of an image with other images, text, and semantic labels. Furthering the OCR work done in the Shakespeare Quartos Archive Project.
Setting of Quarterly Benchmarks	Initial benchmarks stored in Basecamp, and adjusted during the research and development phase of year one.
Deployment of Project communication tools (wiki and skype)	Early project communications used Basecamp as a way of sharing documents among project collaborators. Tools were replaced by a looser combination of using email, Skype at weekly meetings, and email to share feedback on features being developed.
Mid-year Worksprint	Discussion of the tool as it stood. Decision to change the CSS and functionality of semantic linking and OCR.
Initial Review of Tools	Same as the Mid-year Worksprint
Alpha Version of Tools to all project partners	February 2011

Year Two Goals:	Actual Accomplishments:
Year Two Meeting	Identification of TILE tools for integration (May 2010)
Software Testing and Refinement of Alpha Version	Refinement of Alpha version to include improved Scalar Vector Graphics (SVG) image tagging tool and Auto-Line Recognition tool (ALR). The TILE interface, of which two draft interfaces were designed during the first year, were implemented during these months and began to establish the user interface

	connections between the toolset.
Indiana University Worksprint I	Developed specifications for the text labeler, discussed documentation, and collaborated to code and write (Sept. 2010)
Additional Development	Established two additional ways to prioritize and respond to feedback about the tool: a discussion forum on the TILE website where questions from partners could be asked, and a software ticketing system, where bugs could be organized into milestones. The focus was the release of TILE 0.5, an internal alpha version of the software, which was completed October 31, 2010 and shared with partners. Feedback on this initial release was provided by partners, and identified a variety of bugs in the software (some known, and others previously unidentified), as well as desire for interface changes.
Alpha Refinements	Refinement of tools focused on a public release of TILE, as well as focus on abstracting the software into a larger framework that was more modular and reusable in the future. The February 14, 2011 release of TILE 0.9 was the first public release of the software, and included a redesigned website, documentation aimed at both end-users, and a "sandbox" where new users could test-drive the tool without downloading the software.
Indiana Worksprint II (new project activity)	The 0.9 release was followed by a March worksprint in Indiana, which included further refinements with a focus on the development of import and export tools in TILE to support a variety of data formats.
TILE 1.0 Release	Released July 21, 2011

A key request of our partners was to establish a larger content management framework for image-based editions on the web. This same unmet need was written in the original proposal, which noted a history of annotation tools within digital humanities that have never achieved adoption or never hit a critical mass of use to be sustainable in the long-term. The TILE "environment" or "framework" needed a way for the set of tools to seamlessly work together, however while the tools and features of TILE itself were still actively under

development, it was difficult to achieve this larger integration. As a result, the development of the larger TILE framework was held until the second year of the grant.

A decision early in the grant to focus on research and development during year one (investigating and prototyping tools for zooming into images, and developing a reusable SVG image tagger) put considerable pressure on the software development team during year two to quickly produce a stable version of software, as well as react to the feedback of testing partners in a timely fashion. At the conclusion of year one, the TILE team had made significant progress in developing tools, but these tools weren't connected in a single environment or easily shared with partners.

The TILE team underestimated the challenge of connecting these tools, which took far longer than anticipated. By adopting an agile and iterative approach to our software development, we conducted weekly TILE meetings where the MITH development team would respond to testing that was based out of Indiana. When tools were first integrated, it was common for a single bug fix to cause another bug, so this continuous feedback loop was invaluable for developing solutions.

As feature requests and bug reports were made by our reviewers and users, the development team was forced to make changes to the software that sometimes slowed down the overall development. In order to prevent such changes from affecting the overall development, significant effort was put into making the software modular and abstract. This up-front investment proved to be very useful, however initially requiring significant focused energy. As a result, several features of the software, including fully-featured import and export functionality, were not implemented until the final months of the grant.

Two specific challenges were overcome in the successful release of TILE:
- The creation of a javascript-based software framework that links tools within TILE's environment, establishing a modular architecture where plugins in TILE can pass data to other plugins, manipulate user interface components, and add functionality to the environment.
- The development of a semi-automated auto-line recognizer (ALR) in javascript, with accuracy that far exceeds earlier development and may be reusable in different contexts in the future. The grant proposal noted that all existing tools are manual and depend on a human user to create relationships between the image and text.

Audiences
Two public releases of TILE were made during the months of February and July 2011, and drew attention from the Digital Humanities community including hundreds of tweets, being featured on "Digital Humanities Now", and eliciting feedback on the official TILE forums as well as direct email contact to members of the TILE team about its possible use. These releases included the software, as well as documentation directed at end-users and developers.

While TILE was received positively by the Digital Humanities community, it does not appear to be commonly used by the scholarly community. The grant proposal noted that no similar tools have achieved this, which may be less of a technological issue and more of a social problem. Given the relatively brief time since its public release, there are opportunities for this to change as additional people are trained in how to use the software.

Evaluation
A series of formal and informal evaluations were conducted throughout the grant. During

the first year, evaluative research was done to determine what emerging technologies and approaches could be adopted by TILE, and in the second year usability research was done on the TILE environment.

Early in TILE's development, key Javascript libraries were identified as necessary components of the software's architecture: jQuery, and Raphael. In an effort to not "recreate the wheel", the TILE team proceeded to conduct technical evaluation of these libraries and engaged in a series of prototyping efforts that identified where custom code needed to be written, and where existing code could be adopted. The evaluations of software developers were done in conjunction with the TILE partners, who offered specifications for tools during the first year in order to scope the functionality of individual tools. This evaluation and subsequent software development provided building blocks, including TILE's image markup tool, reused later in the project's development.

As the TILE environment took shape and connections between tools were visible at an interface-level, usability testing was conducted. Prior to each weekly team meeting, two graduate students at Indiana who were staffed on the project helped to identify bugs and questions related to the changes made in the software. The week-by-week evaluation of the tool provided granular feedback, while larger usability testing was conducted by select project partners and with graduate-level classes. Feedback from Ginny White at the University of Oregon provided insight into how the workflow fits into the way they currently process transcript and image-linking. User testing done by John Walsh and Julianne Nyhan in their classes provided feedback on how groups interacted with and understood TILE, which informed a larger number of interface changes between TILE 0.9 and 1.0.

Outside of planned evaluations of the software, we received feedback and ideas through the TILE website and forums. While much of this feedback was centered around technical issues and questions related to the software's functionality, it shed light into the ways that people wished to use the software. For example, it was noted by David Lee Miller of the University of South Carolina that TILE's image annotation works for flat digitized imagery, but its use is limited when page images are warped or distorted. We also received feedback on the overall presentation of TILE's functionality from a graduate fellow in the UVa Scholars Lab Praxis Program, who thoughtfully blogged about our website.

Continuation of the Project and Long Term Impact
At the conclusion of the TILE grant, the open source software has been made available for download on Github and the MITH website, along with documentation written for both developers and end-users. There remain questions about the exact form TILE will have in the future, based upon the challenges of establishing use within the community and the need for funding additional development and support. It is also clear that lessons learned through the development of TILE will be, and have already been, incorporated into other Digital Humanities projects.

It was noted in the grant proposal that the software would be "made sustainable through use and continued development in the community." We still believe this to be true, however the reality of funding and staffing changes have changed the open source examples with which we may compare TILE. At this time, TILE exists as an open source project in a code repository where the community is encouraged to fork and contribute. This is different than a Digital Humanities tool project such as Zotero or Omeka that have dedicated software developers and support. TILE users will continue to be able to download the product of this grant via Github

and the MITH website, however our forums have been closed due to lack of direct funding for the project and this type of user support.

Despite dedicated development to TILE itself, the architecture that was developed as part of the grant has significantly influenced ongoing software development at MITH and initiatives that the team is involved in. A version of the modular architectural model that was developed for TILE is being tested as part of MITH's involvement in the Open Annotation Collaboration, and the improved SVG image tagging tool has been integrated into various projects. The focus of developing TILE in a modular fashion means that individual components can easily be used in different contexts, which was intentionally done to make the software are reusable in the future as possible. The spin-off of software is fairly common, especially within open source circles and the digital humanities community. TILE was originally conceived as the successor to AXE, an image-annotation tool that was funded by an NEH Level-One Startup Grant. The development of TILE greatly benefited from lessons learned from AXE, particularly those related to the limitations of image annotation within the web browser.

Grant Products
Information about TILE, including documentation written for end-users and developers is available on the TILE website: http://mith.umd.edu/tile/. Code developed as part of the TILE project is available for download from Github: https://github.com/umd-mith/TILE

During year two of the grant, TILE was discussed and presented at the following conferences, workshops, and meetings. Presentation materials are available by request:
- 2010 Interedition Bootcamp
- July 2010, Geoffrey Rockwell Retreat on Text Analysis/DH Tools
- 2011 CHAIN meeting in Oxford
- February 2011, MITH API Workshop
- June 2011, Digital Humanities 2011 Poster Session