Catalyst Fund Proposal

Proposal Title:	Born-digital Content Viewer for the ArcLight Project
ID:	12
Institution:	Emory University Libraries
Requestor:	Rosalyn Metz, Director Library Technology and Digital Strategies
Budget:	$35,000
Goal (as pulled from the application):	"This project seeks to contribute a born-digital viewer to the ArcLight project. This viewer would allow users to view and browse folder structures and their contents as might be seen on a computer. In addition to providing a viewer to the Libraries Archives and Museums community, the project will also provide proof-of-concept for integrating other viewers into the ArcLight applications."
Description:	This project seeks to contribute a born-digital viewer to the ArcLight project. This viewer would allow users to view and browse folder structures and their contents as might be seen on a computer. In addition to providing a viewer to the Libraries Archives and Museums community, the project will also provide proof-of-concept for integrating other viewers into the ArcLight applications. The total budget for the project is $138,522.50, Emory University Libraries is requesting a minimum of $35,000 and will provide any funding beyond the monies provided by the Catalyst Fund.
	Archives and collecting repositories are collecting more born-digital archival material, and archives still struggle to provide meaningful, contextualized access to these files. We propose to build an application that will present end users of born-digital archives with the ability to view and navigate this material. The application will not only provide access to file content, but also will preserve and display the original file structures of born-digital archival material and will mimic the navigational experience of using the original media.
	We intend to build this application with the preservation of a born-digital file's original characteristics in mind: the file's original metadata will be maintained, either by harvesting the metadata during normalization or by using file viewing software capable of rendering a wide range of current and legacy file formats. Furthermore, enhanced search and discovery functionality will greatly contribute to the overall quality of the user experience. The result will be a system able to accommodate the unique characteristics of born-digital archival material and leverage their research value.
	The proposed born-digital content viewer will be shared and usable by members of the LAM community that hold and are committed to preservation of born-digital archives.
	Work Plan for the Project 1. Gather requirements for a viewer that will meet the needs of users of born-digital archives 2. Develop a viewer based on the requirements gathered at the beginning of this project

| | 3. Contract with LYRASIS for ArchivesSpace hosting
4. Migrate metadata for our born-digital collections into ArchivesSpace
5. Implement ArcLight
6. Integrate viewer into ArcLight

Resources for the Project

Born-digital Archives Fellow: The Born-digital Archives Fellow will work on gathering requirements for a viewer that will meet the needs of archives that collect born-digital materials and disk images. The fellow will work with not only Emory staff but also with the ArcLight project to determine users requirements for this type of viewer. The salary and fringe benefit costs for a Born-digital Archives Fellow are included in the overall budget.

Hosting of ArchivesSpace with LYRASIS: ArcLight brings together Blacklight and ArchivesSpace to support discovery (and digital delivery) of information in archives. As such, an implementation of ArchivesSpace is a necessity to begin the development of the application. While Emory currently plans to implement ArchivesSpace on a larger scale, they do not currently have a working version of ArchivesSpace. An estimate of the cost for 1 year of hosting ArchivesSpace with LYRASIS has been included in the overall budget.

2/3 of a Software Engineer s time: Once requirements have been gathered for the viewer, the developer will implement ArcLight at Emory, develop a viewer based on the requirements gathered, and integrate that viewer into ArcLight. Two-thirds of the salary and corresponding fringe benefit costs for a Software Engineer's time are included in the overall budget.

Results of the project will be shared with the LYRASIS community through a white paper presenting our project results, a presentation at the LYRASIS member summit, collaboration with the Stanford University-led team that develops ArcLight, and code contributions to that project.

The ArcLight product is open source and freely available to the LYRASIS community, as well as the LAM community more broadly. All code will be deposited into GitHub and licensed under an open source, creative commons, or similar license. |
| **Comments from Field Reviewers:** | 1. A "born digital" viewer for ArcLight is an extremely exciting idea that has the potential for broad uptake within the world of archives & research libraries. My key uncertainty about this proposal is its potential intersection / overlap with projects already scoped out or in development within the BlackLight / ArchivesSpace contexts. The fit with LYRASIS is excellent, however.

(continued next page) |

	2. Considering that Emory does not yet have a working version of ArchivesSpace, and Arclight has not yet been released (although a "minimum product" is scheduled for Spring/Summer 2017), I think the budget for this project is rather optimistic. The proposal also does not state if Emory is using or is familiar with any other Hydra/Fedora Commons projects, which can be a considerable learning curve. However, their proposal for developing a viewer for using the Arclight discover tool is admirable. If Emory is indeed willing to continue to fund the project, it would be really great if they were able to develop something. Born-digital objects have been and are very hard to manage, and archives and libraries certainly need better tools to deal with this increasing portion of their collections.
	3. I understand that space is short, but this proposal assumed a lot of a priori knowledge about the problem space. It's a well-described and defined solution, but it's unclear how it fits with the current landscape and how organizations are currently hampered without it.