

**The Media Ecology Project:
Online Access for Scholars Adds Value to Media Archives**

OR2015 Conference Proposal

Mark Williams (Dartmouth), John Bell (Dartmouth), Mark Cooper (U South Carolina)

The Media Ecology Project (MEP) is a coalition of scholars, archivists, and technologists dedicated to expanding the scope of interaction between the academy and the archive. Directed by Prof. Mark Williams at Dartmouth College, MEP enables new forms of digital access to and scholarly analysis of moving image collections and visual culture more generally. The scope of MEP's work toward this goal includes exploring new methods of critical human and computational analysis of media, developing networks between institutions that expose existing archival collections to new audiences, and building tools that facilitate automated sharing of rich cultural data and metadata among software platforms.

MEP is designed to promote efficient cooperation and produce motivated engagement with cultural memory artifacts by academic and scholarly communities. We support close textual studies of the subject matter, production, reception, and representational practices of media. In doing so, MEP also seeks to advance fields of scholarship surrounding these materials and promote a greater understanding of the development and impact of historical media. Raising awareness of these important historical collections is the first step to protecting and sustaining them.

MEP has engaged a wide variety of individuals and institutions to develop a network of stakeholders committed to working to advance its goals. They include:

Archives—The Library of Congress, WGBH, Moving Image Research Collections (MIRC) at the University of South Carolina, UCLA Film & Television Archive, American Archive of Public Broadcasting, Peabody Awards Collection (University of Georgia), UCLA NewsScape Library, and Films Division of India (Mumbai, India).

Scholars—DOMITOR scholars team, lead by Prof. Tami Williams (The University of Wisconsin, Milwaukee) and featuring Prof. Frank Kessler (Utrecht University) and Prof. Laura Horak (Carleton University); News Materials scholars team, including Prof. Mark Cooper at The University of South Carolina; The Red Hen Lab, lead by Prof. Francis Steen (UCLA) and Prof. Mark Turner (Case Western Reserve University); Gender Research Institute (GRID) at Dartmouth.

Developers—The Columbia Center for New Media Teaching and Learning (CCNMTL), The Alliance for Networking Visual Culture (ANVC), The Virtual Environments and Multimodal Interactions Laboratory (VEMI), Still Water for Network Art and Culture (UMaine); Center for Computing and Visualization (Brown University).

MEP has designed, modified, or produced several software tools to facilitate the creation and exchange of time-based media annotations with the goal of automating scholarly contributions to archival collection management systems. These tools include:

MediaThread—A classroom platform developed at Columbia University that has been extended to support publication of time-based annotation metadata and integration of external controlled vocabularies for tagging. Release of the MEP-compatible version of MediaThread is confirmed for February 2015.

Scalar—A digital publishing platform developed at The University of Southern California that is being expanded to support import of time-based annotations and controlled vocabularies for tagging. Release of the MEP-compatible version of Scalar is expected in Q2 2015.

onomy.org—A new tool developed by MEP that facilitates collaborative creation and sharing of controlled vocabularies by small research groups. The initial version of onomy.org was released by MEP in 2013. An upgraded version is set for deployment in early 2015.

MEP Metadata Server—A new tool under development by MEP that aggregates and distributes time-based media annotation metadata among workspaces (current MediaThread and Scalar) and archives. The initial version supports the W3C standard for open annotations on the web and is slated for release in 2015.

The goal of building these tools is to create a software ecology allowing analysis created in one piece of software to be shared in many. For example, one workflow might see a group of researchers create a vocabulary in onomy. They could then use the terms they defined to tag a Library of Congress media file with annotations using MediaThread. Those annotations would be registered with the MEP Metadata Server, which would in turn make this metadata available for publication in a Scalar book or for harvesting by the Library of Congress to enhance their own records about their collection.

There are four pilot projects currently in development for MEP:

Paper Print Collection—The historically significant Paper Print collection at The Library of Congress is the equivalent of the Rosetta Stone for those who study moving image history in relation to visual culture. In conjunction with Prof. Tami Williams at The University of Wisconsin, Milwaukee we have enlisted several members of the renowned DOMITOR research society to study these early silent films that are being digitally transferred and preserved at the Library of Congress. The scholars are coordinated into five Special Interest Groups to collaborate on examining varied aspects of the films in the collection. Presentations of their findings will begin in Spring, 2015.

In the Life—*In the Life* is a historic public television program that assayed the history of gay and lesbian lived experience in the United States. The entire run of that program, plus all of the associated materials involved in the production of that program, will be preserved and placed online in Spring, 2015 by the UCLA Film and Television Archive. Scholarly participation will include prominent members of The Society of Cinema and Media Studies and also the Gender Research Institute at Dartmouth.

Historical News Media—Scholars from multiple institutions are using MEP-connected tools to design a broad study of historic newsreels, newsfilm, and newscasts housed in disparate collections including Open Vault at WGBH in Boston, The Moving Image Research Collections at The University of South Carolina, The American Archive of Public Broadcasting, the Peabody Awards Collection at The University of Georgia

Special Collections Libraries, the UCLA Film and Television Archive, and The Library of Congress. Participating scholars include Prof. Mark Cooper at The University of South Carolina and Prof. Ross Melnick at the University of California at Santa Barbara. Along with MEP director Prof. Mark Williams at Dartmouth College and Prof. Sara Beth Levavy at The University of Utah these scholars will co-edit a compendium on news media history in the U.S. to the 1990s and early 2000s.

Films Division of India—This pilot is focused on studying the legacy of documentaries and informational films archived at Films Division in Mumbai, India. Films Division has produced state-sponsored documentary, informational, and experimental cinema since India's independence in 1947. This institution is also working to create a national archive that will curate and study the history of Indian cinema overall. This pilot study was initiated in December, 2014 as a result of MEP conference participation and archive visits in New Delhi and Mumbai. An international team of scholars has already initiated the study of these materials.

In addition to work already completed and underway, MEP has been exploring new initiatives that support its overarching mission of developing new forms of archives-based scholarship. Principle among them are two lines of inquiry that are being shaped into discrete research and development projects:

Scholarly Access Network—Many archived films that would be of great interest to the scholarly community are currently held in collections that are either not publicly available or available only at the physical location of the archive. There are several obstacles, both legal and technological, to providing networked access to these materials so that scholars using the MEP toolset may study them. In conjunction with our archive partners we will pursue white paper recommendation for a legal framework under which archives could allow restricted network access to materials encumbered by intellectual property concerns. Once the legal framework is available, we will pursue a technological implementation that would support secure scholarly access to IP-restricted archival materials over open networks.

Machine Annotation—Building on the existing work of several MEP participants, this initiative would examine how to create machine-generated annotations and include those annotations in the MEP metadata exchange ecosystem. Projects like ACTION (Dartmouth) and NewsScape (UCLA) have developed technology that algorithmically analyzes moving image files and generates metadata that describes various aspects of moving image texts, such as how objects move across the frame or how on-screen text corresponds to the action being shown or dialogue. These automatable tasks provide a new and exciting methodology for distant reading of large volumes of films, but currently there is no easy way to correlate metadata resulting from running different types of analysis on the same film. MEP's Metadata Server provides a potential solution to this problem but would need to be extended to support new types of metadata to connect machine vision analysis tools in the same way it connects human-generated annotation tools like MediaThread and Scalar. This line of inquiry would develop a common metadata standard that describes the complex relationships at the core of machine vision analyses and then extend the MEP Metadata server to support high volume, automated metadata exchange between machine vision tools.

MEP will enhance the functional discoverability of materials across multiple archives and augment future efforts to produce new forms of digital scholarship about these archival materials. The MEP archival connections are being built on public standards like the Open Archive Initiative and the W3C Open Annotation format. Use of these widely available standards is key to realizing an ecology of applications that encourages bidirectional communication and shares information as peers, working with archives as not only key sources of raw materials but also as advocates for and consumers of new analysis and scholarship.

Panelists:

Mark Williams will present an overview of the project, including an introduction to the relevant platforms and tools engaged (Mediathread, Scalar, Onomy), and the four pilot projects (Paper Print pilot with The Library of Congress and DOMITOR; “In the Life” pilot with UCLA Film and Television Archive; Historical News Materials pilot with various archives, including WGBH, MIRC at The University of South Carolina, and American Archive of Public Broadcasting); Films Division of India pilot with assorted international scholars and institutional support within India).

John Bell will explain the technical architecture behind MEP’s third party federated metadata network and how it connects repositories to client applications like Mediathread and Scalar. The MEP network is built on core standards including the W3C Open Annotation format, FOAF, and RDF. MEP has chosen its mix of technologies based on a belief that a certain set of shared functionalities is critical to the successful use of a federated archive structure. Standards for access are of course critical, but functional integration requires systems that share a baseline understanding of concepts like provenance, rights management, annotation, and vocabulary while allowing individual member archives to maintain their own specific metadata ontologies that extend beyond those base concepts.

Mark Cooper will discuss how MEP can enhance classroom teaching and scholarly research that makes use of the Islandora Repository of the University of South Carolina’s Moving Image Research Collections (MIRC). Specifically, he will consider a project to familiarize students with the conventions of archival moving images through the process of annotating them and another collaborative project to use the MEP VRE to accumulate descriptive metadata for theatrical and television newsfilm coverage of protests.