# Proposed topics

**See also Proposed topics by archival function.**

## Overview

In addition to a 5 minute lightning talk we also require everybody to **propose a topic**. This does **not** necessarily mean that you are an expert or that you are willing to lead a discussion on this topic. Proposed topics could entail:

- **topics you would like to discuss in more detail with the group** - eg a problem (or a solution) you have encountered
- an **aspect of your work** that you would like to share with the group (please bring examples of workflow, documentation etc with you)
- a topic they would like to **learn more** about

After the deadline we will then turn this into a list and ask people to indicate their preferences/interests ahead of the event and use this to shape the discussions.

In each please explain your level of knowledge or experience and comment on the proposals. Please use the following format:

**Short phrase summary of topic**

Name, Affiliation

Paragraph-long statement describing topic

## Proposals



### SIP Creation and Rubymatica

Tom Laudeman, University of Virginia (AIMS)

I would like to demonstrate Rubymatica and discuss issues related to processing of files to make them suitable for a Submission Information Package (SIP). I am the author of Rubymatica. It is a Ruby port of the SIP creation phase of Archivematica. My efforts were supported by the AIMS team, especially the digital archivists. (Many thanks to the Archivematica people for blazing the trail.) A key question is: How much processing should we do before initial assessment? The demo is just a few minutes, but I hope the discussion will be lively.

### Describing Electronic Records: What's Useful to the Researcher?

Mark Matienzo, Yale University (AIMS)

I'd like to discuss what people think is useful and practical information to incorporate about electronic records into archival description. A good example of this is extent - is extent in terms of data size (e.g. megabytes/gigabytes) sufficient and appropriate? Are file/directory counts useful to researchers, or are they potentially misleading? I'd also be curious how people are approaching description of records like websites as well - whether they're describing sections of a site's information architecture, and what other information would be crucial to provide in a access system for archival description.

### What we can do to help researchers to work on email archives?

Peter Chan, Stanford University (AIMS)

I would like to talk about "Muse" which is a program for browsing long-term email archives using data mining and visualization techniques. The program is created by Sudheendra Hangal from the Mobisocial laboratory at Stanford University. (http://mobisocial.stanford.edu/muse/)

### How informed are our depositors?

Simon Wilson, Hull University (AIMS)

I would like to discuss the topic of how informed are depositors are (or are not) about born-digital archives and the critical difference regarding their expectations and assumptions relating to the preservation and access of born-digital material compared to traditional paper archives. What do they expect of us? How can we ensure that their expectations are reasonable?

### Collaboration and Institutions?

Gretchen Gueguen, University of Virginia (AIMS)

I would like to talk about how institutions can effectively collaborate in the creation/management of born-digital archives. Each institution has it's own quirks and local practices, but to be really effective we need to figure out how to share the burden. In a larger sense we could also discuss what such collaboration would really mean, just sharing information? infrastructure? development of tools? Collaborations take effort, so what would make them worthwhile?

### SIP formation procedures - What needs to happen before our systems "Receive SIP"?

Courtney C. Mumma, City of Vancouver Archives (Archivematica developers)

What work needs to be done before a donor transfer is ready for processing? Archivematica is on the verge of expanding to include SIP-preparation procedures prepending the processes currently packaged in release 0.7. A certain amount of analysis and configuration needs to occur before the SIP is in a state that it can be processed and transformed into an AIP. At the very least, formation of the SIP from a donation requires accessioning and administrative tasks, forensic imaging and analysis, log creation, metadata creation, malware checking, partial arrangement, appraisal, identification and unlocking of encryption, identification of confidential or restricted information, assignment of identifiers, etc. We've outlined a number of draft requirements on our project wiki (http://artefactual.com/wiki/index.php?title=SIP_Creation ) and you can see Artefactual's plans for Archivematica on their development roadmap (http://archivematica.org/wiki/index.php?title=Development_roadmap ), but I think it would be useful to compare notes with others who have "pre-Ingest" workflows and experiment with tools they are using.

### Examples of Patron Use

Matthew Kirschenbaum, University of Maryland

I would like to discuss known examples of patrons working with materials in born-digital collections. What case studies are already out there? What examples can we point to of born-digital materials that have been consulted and cited by scholars in their work? What wisdom is available anecdotally from those scholars who have worked with born-digital materials, either online or by visiting a collection? What provisions have institutions made for researcher access? How can researchers be educated about both the potential of born-digital materials, and the kind of problems and challenges they should expect to encounter?

### Trust: Perceptions and Demonstrations

Ed Fay, London School of Economics

Trusted repositories have been discussed for many years, and attributes proposed based on OAIS (TRAC, nestor et al.) with current work taking the proposed certifications towards ISO standards. I would like to discuss whether these demonstrations meet the requirements or expectations of some or all of our user constituencies (technologists/system builders, depositors, curators, end-users, organisational management) - in essence whether the demonstrations match the perceptions. A further aspect which I think is worthy of discussion is how digital trust relates to traditional trust. Curating digital material is only an extension or format shift of a mission which organisations have been carrying out for many years. Rarely do depositors require demonstrations of strong-room standards-compliance (e.g. BS5454) so is trusted repository status an exercise purely for the specialists? Will trust continue to be placed in the organisation regardless of technical demonstrations? What, if anything, do trusted certifications need to demonstrate that they are not already?

### Capturing People’s Digital Contexts: original order, appraisal and description.

Catherine Hobbs, Library and Archives Canada

Creators of archives are living digital or hybrid lives in terms of the media of their records and they are also living with portable devices. Archivists have traditionally served to capture the context of record-keeping and this can have very specific shades within the lives of individuals.

I would like to discuss how we interpret original order and what we ask creators about their record-creating habits and decision-making with the digital realm (e.g. Perhaps we shouldn't assume the same priority for records that are digitally abandoned but yet kept somewhere on a hard drive). How can we glean the creator’s perspective on their own record-keeping and life with technology during the site visit and appraisal, and then transfer these elements to description? We can’t rely on automatically harvested metadata or even emulation to capture these aspects for us. New approaches may involve asking the right questions, changing appraisal values, expanding description or even modelling relationships between technologies and documents in ways that make apparent the personal context of these records.

### Training our donor/creators

dave thompson, Wellcome Library

I would like to discuss the problems that we have in engaging donor/creators to transfer born digital material to our library.  Many of our existing donors create material digitally, but continue to prefer to print it for transfer to us. Many are organisations, we do have individual donors.  Many have been passing their material (Physical)  to us for ages.  We have spoken with our donor/creator community about the transfer of digital material, in principle most see the advantages and opportunities.  But few do it. Somehow it is too difficult.  We provide information on the digital transfer process, we explain how we'd like the process to work. We meet to discuss the means of transfer, we provide portable hard drives. We explain that we don't scatter their material willy nilly across the internet for all to see.  It's difficult to know what else to do!  I'd be interested in discussing the expereinces of others.  Is engagement with donor /creators and their digital material much more time consuming?   Are donor/crators much more 'possessive' of their digital work? Is it that they haven't managed to organise their material digital in the same way that they have their physical? Is there a perception that digital material is somehow more ephemeral and less 'important'?

## Ethical issues surrounding acquisition and access to e-MSS

Helen Broderick, British Library

I am interested in discussing ethical and privacy issues attached to born-digital archives and in particular whether any new or different issues are raised by this different type of record. For example do different privacy considerations need to be made for emails that are not necessary for paper correspondence?  Should all email addresses be redacted as a rule? Are researchers any more likely to email an author than they are to visit their home if both addresses can be found in their archive?

Do these new challenges mean that archivists will increasingly be called upon to learn new skills or work more closely with digital archivists or IT departments? Whilst time restrictions on paper archives can be easily administered by archivists the situation is more complex when restricting material within born-digital collections.

## Managing "Restricted Access"

Seth Shaw, Duke University

How do we identify, preserve, describe, and provide access materials to "restricted materials?" Many times donors give us material with very sensitive content with varying degrees of awareness. Software is available both to uncover deleted records but also to identify PII (personally identifiable information) of donor and others but these usually miss some things. Do we place spacial, temporal, or group based access restrictions on the materials we can identify as restricted? How do we manage access to originals, redactions, and/or derivative access copies?

## Using Forensic Software to assign Metadata to Born Digital Archives

Peter Chan, Stanford University (AIMS)

The heterogeneous file formats and content types in most born digital archives post a serious challenge to archivists in assigning metadata to individual digital objects. Stanford University Libraries has been exploring the use of forensic software since April 2010 to generate technical metadata and to assign descriptive and administrative metadata to several born digital collections. The technical metadata includes checksum, file format, file size, file creation date, last modification date, and last accession date; descriptive metadata includes archival context (series, subseries, etc.), subject headings, and source media, and administrative metadata includes primarily access restrictions. I would like to share this experience and to receive feedback from other people.

## Discussing Digital Issues with our Stakeholders

Alison Hinderliter, Newberry Library

As cultural institutions continue to receive collections, and as the collections become more digital in nature, the process of accessioning and delivery of documents, which used to be within the purview of the curator, archivist, and reading room staff, has now become the concern also of library administrators, the IT department, and the finance department.  How should the conversation begin with departments that have never before had to deal with the questions of long-term preservation and storage of archival materials in varieties of formats?  As a related issue: What can be expected of the smaller institutions with smaller staff and budgets, and does collaboration with other institutions include the referral of larger collections to larger, more well-funded institutions that are better equipped to handle the complexities of preserving and providing access to these collections?

## Establishing Digital Archives as "the Norm" in Archival Practice

Erin O'Meara, University of North Carolina

This is somewhat complementary to Alison's topic above. I want to talk about how we fully realize born-digital materials as normal components of a special collections and/or archives workflows. Many repositories are hiring specialized staff for born-digital. I want to ask the group and discuss how we move to digital as the norm within our competencies, workflows and policies. Part of my strategy in a large, multi-special collection library has been to empower archivists (curators, technical services and public services staff). We have designed basic procedures and user-friendly tools so they can begin to feel comfortable with handling born-digital materials. I would like to hear what other repositories are doing to engage archivists in this area.

## Questions of scale and forward planning

Susan Thomas, Bodleian Library, University of Oxford

As born-digital archives become business as usual in our institutions how do we ensure that our approaches and workflows are appropriate for all of our collections, regardless of their scale or import? I'm specifically interested in how this applies to accession, appraisal, arrangement, sensitivity review and description. I'd also like to get a sense of how institutions are planning processing projects involving born-digital archives - how much time is factored in for the 'digital activities' and how is that being costed? What do we need to know to plan a project accurately and how feasible is it?

## Advocacy & Staffing

Michael Forstrom, Beinecke Library, Yale University (AIMS)

I’m interested in discussing staffing or program models for managing born-digital material in collecting repositories. As a group I imagine we can identify some variation in the following: types of institutions, job titles and responsibilities, places in our institutional hierarchies, skill sets and experience, resources (e.g. staff, infrastructure, financial, educational) at our disposal, and so on. Is there a natural/common evolution to our program development? Media shows up in collections, accumulates, gets the attention of staff, becomes a part-time responsibility of one staff member... and for how long, until there’s a mandate, dedicated expert staff, adequate resources?

### Assessing User Needs & the User Experience

Erika Farr, MARBL, Emory University

I’m interested in discussing how we can best (or better?) assess the needs researchers have for born-digital and hybrid collections. In keeping with this interest in researcher needs, I would also like to discuss different approaches to monitoring the researcher experience as users interact with born-digital and hybrid archives. While we are pursuing user feedback at Emory, I am not sure we have discovered the most efficient or effective means and I am anxious to learn how others are approaching this task. Questions for this topic could include: How can we assess both current and future researcher needs? What sustainable and extensible approaches can we take to soliciting feedback from users? How can we use researcher feedback to improve our tools and points of access?

This topic shares a lot common ground with Matt's topic on patron use.

### Rights Management / Access Control

Brad Westbrook

The more complex and heterogeneous a repository’s collection, the more difficult it is to manage the rights of and access to the contents. How are repositories capturing / expressing the rights status for the diverse objects in their collections? What procedures are being used for updating the rights status of objects, as they change over time? And, in echoing Seth’s topic proposal, how is the rights status information being used to manage access? It would be illuminating, I think, to see the range of practice and processes used to address these matters.

### Acquisition Strategies

Melissa Watterworth Batt, Dodd Research Center, University of Connecticut

I am interested in discussing new strategies for acquiring born-digital materials and new methods of reflecting and capturing donor relationships in agreements more precisely. In my experience with creators in a hybrid environment for a few years, my curatorial practice has changed considerably. There is some precedence for this with personal papers in the paper world, however what used to be the acquisition process is now the acquisition project. I find my interactions with donors are happening with increased frequency, much earlier in their career, and that these 'agreements' are negotiated at a rate, practically, with each accession. How can this process be managed and be successfully scaled? Is it possible to implement a number of acquisition methods and sustain them?

### Models for Access

Dawn Schmitz, University of California, Irvine

I would like to learn more about different practical models for creating access to born-digital materials, including those in hybrid collections. How is access provided in the reading room? Is description of born-digital materials integrated with that of analog materials? How are access policies and procedures, such as those related to making copies, managed? I would be interested to know whether other institutions have experimented with online remote access (whether open or restricted to registered users). While creating greater access, this is not appropriate for every type of collection and it brings about its own set of issues related to: intellectual-property and privacy rights (including those of third parties); scalability; infrastructure; and, in the case of hybrid collections, providing differential levels of access to different forms of materials in the same collection.

### Automatic Tools for Access and Description

Catherine Stollar Peters, University at Albany

I would like to discuss data mining and visualization techniques that can be used to access digital archives. What tools exist/are people using to access information in digital archives in novel ways? How can these tools decrease the amount of manual description required for digital records? Recently, I spoke to a researcher at GE where their preservation strategy is to keep everything and invest in search and retrieval, data mining, and visualization. Can we learn anything from researchers in industry who take a different perspective on digital curation?

### Inter-institutional collaboration and communication

Gabriela Redwine, Harry Ransom Center, The University of Texas at Austin

I'd be interested in hearing more about how different curatorial areas or departments within institutions (e.g., literature, film, photography, art) are communicating about born-digital materials. When digital media arrive, are accessioning, capture, and preservation the responsibility of the department to which the materials belong? Or is there a central person or department that handles everything related to born-digital? What is the current reality and the ideal future?

### Failure and absence

Gabriela Redwine, Harry Ransom Center, The University of Texas at Austin

Failure is central to the work of a digital archivist. Most workflows, procedures, policies, and other sorts of documentation are created with success as the final goal, and presence (rather than absence) as the primary representation. In reality, failure is more prevalent than success, especially for digital archivists working across the legacy - modern media divide. How should we be documenting failure and absence? Which stakeholders might find this type of information valuable? When does failure become success, and when should an absence be described so as to become a presence?

### Change

Aprille McKay, Bentley Historical Library, University of Michigan

Choosing technologies, tools, formats, collaboration partners, workflows and practices requires us to make predictions about the likelihood and pace of change. Project management techniques encourage us to assess the potential threats of changing personnel, policies, and technical and fiscal environments and to create plans to mitigate their effect. What works and doesn't work here? How do we determine when the pace of change is so fast that a wait-and-see approach is appropriate, and when should we push through a plan in the face of uncertainty? What are the peripheral gains attained by pushing on?