

Preservation Needs Assessment for Media-bound Digital Content (MBDC)

Jess Whyte, Digital Asset Librarian, 2018
University of Toronto Libraries, jessica.whyte@utoronto.ca

Problem Statement: Media-bound digital content (MBDC) presents unique risk factors for preservation. Before the University of Toronto Libraries could begin to mitigate those risks, it first needed to assess (1) the extent of these materials within its collections, (2) their locations within the library system, (3) their local collection holders, and (4) the needs of those collection holders. Traditional condition surveys, sampling approaches, or preservation assessment methods were not well-suited to this task because they require a high level of intellectual control for what has often been under-described or inconsistently-described materials. An object-centric approach also leaves out the collection holders, their institutional knowledge, and their understanding of current conditions, workflows, and needs. Below outlines a preservation assessment method for media-bound digital content (floppy disks, CD-ROMs, USB keys, etc.) that addresses these challenges.

Research questions

- How much media-bound digital content do we have?
- Where is it? Who holds it?
- How are those collection holders currently accessioning, stewarding, and/or providing access to that digital content?
- Do those collection holders need help doing any of the above?
- If so, how can we provide that assistance in a way that works?



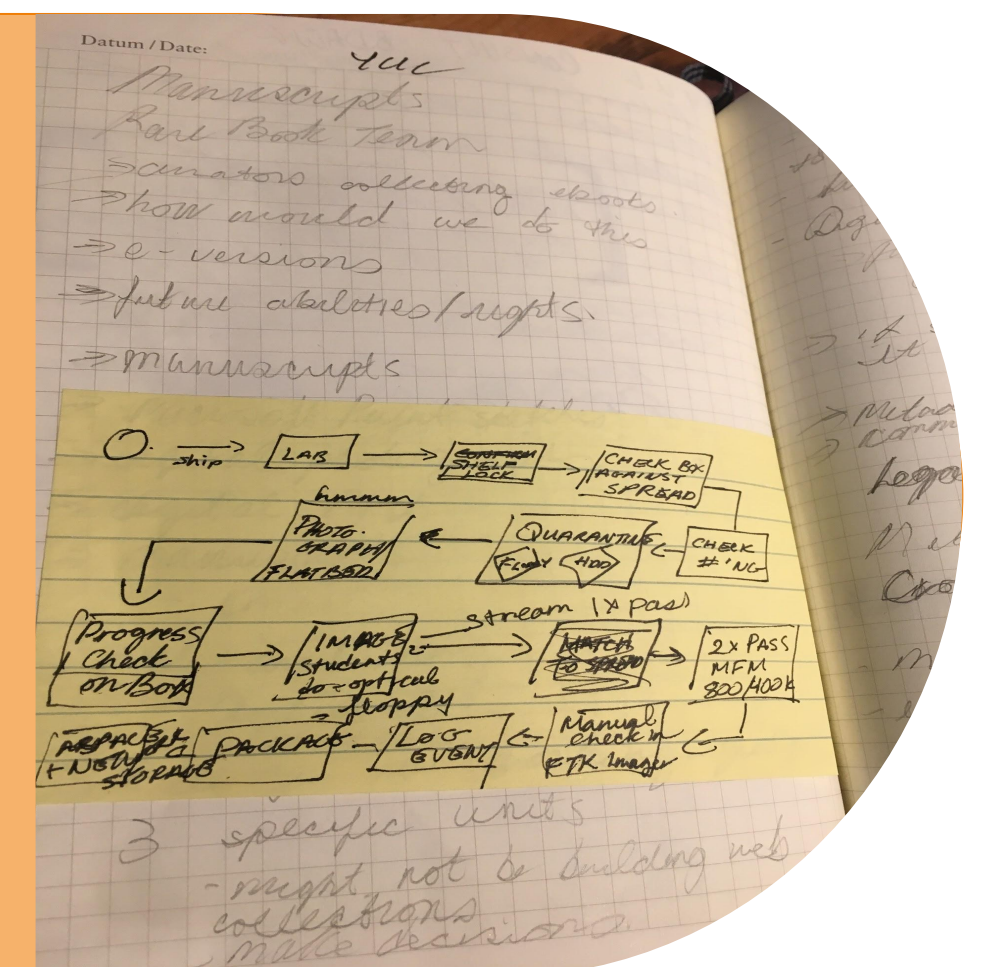
Connect with collection holders

We identified 48 collection holders including representatives from the 44 libraries and archives. We e-mailed or called the identified stakeholders asking them what, if anything, they held within the scope, how they were currently stewarding it, if they would like help doing so, and if they would like to meet to discuss it further. Most had a number of cases, examples +/- or questions and opted to meet with us to discuss it further. **Making direct personal contact was key** here because of the specificity of the scope, time, and our research questions, which were not conducive to a survey or mass-communication approach.



Document Analysis

Document analysis included searching catalog records and Discover Archives for relevant terms (e.g. electronic resource, disk, disc, optical, CD-ROM, USB, floppy, books-on-disk, software, etc.) and conducting refined format searches. However, catalog records and finding aids alone are not ideal as one entry might consist of hundreds of digital media objects. The search is also not physical, and doesn't tap into the institutional knowledge held by staff. Relevant policies and other documentation were also reviewed for history and an understanding of current status or conditions. These included UofT's Governing Council's policy on Information Security and the Protection of Digital Assets (2016), UTL's Guiding Principles for Digital Preservation (2016), relevant minutes and discussions of the Web Archiving working group (2016-) and UTAAG (2015-), current deposit agreements for T-Space and UTARMS, the Library Council's Report on the Task Force on Policies and Procedures for Replacements in UTL Collections (2000), and the Library Council's Brittle Materials Committee Report's recommendations for cataloguing and care of digital items resulting from digitization efforts (1999).



Field visits

Visits typically involved meeting at the library or archive, a semi-structured interview asking what they currently held (probing by providing examples from other collection holders), how or if they provided access, what needs they have, and, if they did want help with media-bound digital content, how they might want that service or support to look. At this point, we discussed their current workflows and systems for stewarding digital content. Collection holders were also asked for a tour of the digital media we were discussing. This gave us a feel for the storage conditions, the extent, the level of intellectual control, and the discoverability of the content. Once site visit notes were complete, they were sent back to collection holders for verification.



Environment Scan

To understand how other institutions were mitigating the risks of media-bound digital content, a document analysis of University of Michigan's Born-Digital Lab, Indiana University's Born-Digital Preservation Lab, the retroTECH Lab at Georgia Tech, New York Public Library's Digital Archives Lab, and the University of Texas Harry Ransom Center's Born-Digital Forensics Lab was conducted. Longer site visits to Yale University Library's Digital Archaeology and Preservation Lab and Stanford University's Born-Digital Forensics Lab were also made.



Analysis

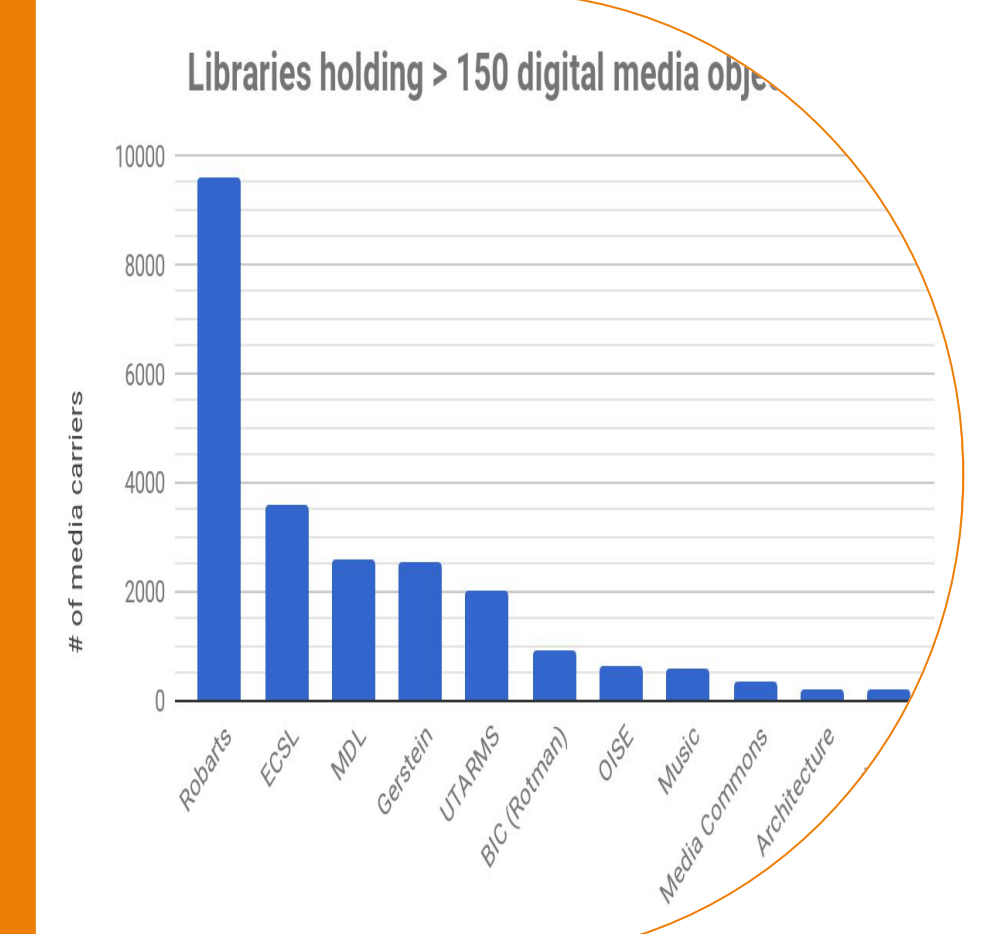
Notes were compiled to come up with holdings estimates, pull out common needs and opportunities for UTL staff, and identify common characteristics or potential solutions from peer institutions.

Finding - extent of MBDC:

Total across UTL: ~25,500 media objects (~75TB),
UTL Central: ~23,000 media objects,
Special collections: ~5000 media objects

Needs Raised By Collection Holders

1. Equipment (hardware and software) capacity
2. Preservation method or workflow
3. Systems and infrastructure (storage)
4. Guidance on collections decisions
5. Determining archival or collection value without intellectual control/ability to review contents.
6. Guidance on copyright implications
7. Guidance on description of digital content and creating links in the catalog and/or finding aids
8. Access methods
9. Staff time



Rec'ns and pilot design

Recommendations:

1. Run a 2-yr pilot with the goal of clearing 25% of MBDC backlog.
2. Provide accessible and secure service point for collection-holders needing assistance.
3. Prioritize collections for format migration.
4. Each participating library or archive designates one named individual to act as liaison for digital content.
5. Provide collection holders with guidance on collections decisions for obsolete media; description of digital content; and a FAQ addressing potential access scenarios or copyright concerns.
6. Work with cataloguing and collections development staff on handling incoming media-bound digital content, particularly supplementary materials.
7. Develop documentation on how to access current and upcoming services (including accepted formats).
8. Digital Preservation Unit offer open hours for staff and researchers.
9. Develop and implement training for collection holders on how to work with service output and other digital content.
10. Develop an Information Risk Management Program (IRMP) for content likely to contain sensitive or personal information (e.g. special collections).
11. Continue to research and refine access methods and reduce mediation required.
12. Develop or adopt a shared vocabulary.

Jess Whyte, Jan 2018
Needs Assessment: Media-Bound Digital Content
University of Toronto Libraries



Overview
A significant portion of the University of Toronto Libraries' (UTL) digital content is media-bound (i.e. stored on physical or electronic media such as floppy disks, optical media, hard drives, etc.). This needs assessment seeks to identify the extent of these holdings, needs of collection holders, and opportunities for long-term preservation and access.