Collaborative Collection Development

Efforts on collaborative collection development continue apace on several fronts. Locally, Columbia continues to engage with its New York City partners, New York University and New York Public Library, in MaRLI (the Manhattan Research Libraries Initiative) whereby researchers at any of the three institutions can apply for access to the partner collections for materials needed for their research. This cross-partner access allows the group to reduce duplication of prospective print acquisitions in select areas by negotiating MaRLI-wide access to electronic content at reduced cost. Stretching our local collaboration into New Jersey, the ReCAP partners (Columbia, New York Public Library, and Princeton) can now borrow from each others off-site collections stored at a shared facility in New Jersey. The next phase of this project will entail addressing the complexities of defining a single shared collection of serials holdings, fine tuning the middleware that enables cross-partner borrowing, and creating a system that will keep track of institutional collection building responsibilities pertaining to the Shared collection. With Ivy Plus, we are engaged with the Web Collecting Initiative (mentioned above) and are exploring a project to analyze the print holdings across the partnership in order to better understand the print corpus underpinning the BorrowDirect partnership. This analysis should allow us to identify opportunities to reduce overlap and expand coverage.

Web Archiving

In the past year the CUL Web Resources Collection Program dramatically expanded its Avery Library Historical Preservation & Urban Planning collection, which now has over 440 websites, as well as continuing to grow the Human Rights and NYC Religions thematic collections and capturing Columbia's institutional web presence. In the spring of 2017 a new "Resistance" collection was begun of websites documenting and embodying grass roots political resistance activities in the United States arising in the wake of the election and inauguration of Donald Trump, including selected web resources focused on: protest/demonstration planning; calls to action and guidelines for political participation by citizens; official statements in response to controversial executive orders; and campaigns to contain and mitigate the effects of specific dramatic policy shifts.
In 2017, Samantha Abrams began serving as the Ivy Plus Libraries Web Resources Collection Librarian; in October, Jean Park began serving as the Program's Bibliographic Assistant. Over its first full year, the Program has dramatically increased the size of its pilot collections: the Contemporary Composers Web Archive went from 51 to 980 websites, and the Collaborative Architecture, Urbanism, and Sustainability Web Archive increased from 130 to 196 websites. The Program also launched its first new collection, the Global Webcomics Web Archive, dedicated to preserving webcomics and creator websites in a wide variety of styles, subjects and themes, in many different languages, created by a diverse group of artists. The Program also has four additional collections in the queue, covering topics such as: geology, politics, banking statistics, and Latin American contemporary art. In its second year, the Program will focus on sustaining its new collections and proposing plans to support increased metadata creation needed for new collections.

Columbia University Libraries has been a member of the International Internet Preservation Consortium since 2011, and in 2017 completed the first year of an elected three-year term on the Consortium's Steering Committee, with Alex Thurman serving as Columbia's representative. Thurman and Abrams additionally served on IIPC working groups and on the Program Committee for the 2018 IIPC web archiving conference in Wellington, NZ.

2CUL Technical Services

Although not exceptionally active, the steering group for the 2CUL Technical Services Strategic Alliance remains as a consultative group for issues of mutual interest. The group wishes Jim LeBlanc well in his retirement. Jesse Koennecke will replace Jim as the Cornell lead on the group.

FAST Five

FAST has been the main area of interest for the 2CUL Technical Services Strategic Alliance and Kate Harcourt reported on progress made by the group in bringing FAST out of research into production at the ALCTS Heads of Cataloging Discussion Group at Midwinter 2018. Since that time, OCLC has committed to bringing FAST into production and has been conducting focus group meetings for product enhancement. Columbia is encouraged by this and has incorporated use of FAST in support staff job descriptions so that more original cataloging can be done to process hidden collections. A barrier to implementation had been that FAST terms do not display in our discovery system even though they are indexed and incorporated into subject facets. Users do not know why they retrieve records and can not follow hot links to other works using FAST terms. With input from cataloging staff, Columbia developed code for Blacklight that displays FAST for records lacking LCSH to better support FAST use in original cataloging. This work may be of interest to other users of Blacklight discovery platforms.

Manuscripts of the Muslim World

Columbia is participating in the Manuscripts of the Muslim World project, a CLIR grant-funded initiative between the University of Pennsylvania, Columbia University, and the Free Library of Philadelphia. The libraries will share an original cataloger (Dr. Kelly Tuttle) who will be responsible for providing original
cataloging and descriptive metadata for manuscripts in Arabic, Persian, and Turkish from the collections of the Free Library of Philadelphia, Columbia University, University of Pennsylvania, Bryn Mawr College, and Haverford College. The grant will provide an opportunity for us to learn how to refine metadata provided by graduate students and subject specialists, use MARCEdit to export brief MARC records to a shared OCLC save file for the project cataloger to supply scripts and other metadata and finally link and distribute high quality descriptive, technical and preservation metadata to local catalogs, Hathi Trust, the Internet Archive, DPLA and OPenn. We may also try to contribute NACO records through Penn’s Arabic Funnel.

**Strategic Planning: Uncovering Hidden Collections**

Columbia continues to make progress on Strategic Directions, especially in surfacing hidden collections. The Hidden Collections Initiative Team reports the following results for 2017: allocation of funding to hire three archivists (two of which are new permanent funding lines from the Provost specifically for hidden collections work), creation of the hidden collection inventory methodology, glossary, data collection tool, metadata standards for all formats and compilation of first annual hidden collections inventory data. The range of materials in the inventory is immense so work this coming year will include prioritizing and scoping projects, examining existing workflows for efficiencies, and refining assessment tools to measure ongoing progress in backlog reduction.

A significant percentage of hidden collections consists of archival materials. Columbia is implementing ArchivesSpace to help establish efficient workflows for exposing legacy collections and processing new acquisitions. Having a single platform to accession, describe and manage archival collections from four archival repositories (Avery, Burke, Rare Books and Manuscripts, and Starr East Asia) will make material available to users more quickly and prevent backlogs from developing in the future. The ability to export descriptions in both MARC and EAD will support discovery and access in our library catalog and local Archives Portal as well as OCLC’s Archive Grid. Challenges include remediating legacy metadata, keeping metadata in various data stores in sync, and creating a process where authority processing and other maintenance is replicated in ArchivesSpace. Our goal is to have the instance populated with legacy data and ready for new data entry by July 2018.

This spring saw a further refinement of the Hidden Collections initiative. The Provost’s Advisory Committee on the Libraries (PACL) asked us to prioritize uninding high risk materials that need to be digitally converted for preservation and user discovery. The faculty recognizes that 20th century analog formats are rapidly deteriorating and that unique historic sound and moving image content is playing an increasingly important role in research, teaching and learning. We are embarking on a one year pilot project to address at risk time-based media (TBM) by establishing workflows and procedures to enable digitization, metadata creation and remediation and preservation at scale. Our goal is to create an action plan to digitize our backlog of 59,000 unique TBM with a focus on high risk unique items of high research value as advised by the faculty. The 2018/2019 project will focus on completing the processing of 4,000 oral histories from the Columbia Center for Oral History Archive (CCOH) and clearing a large backlog of digitized but inaccessible items. Original and Special Materials Cataloging (OSMC) is
recruiting a project metadata librarian to provide original metadata creation and assist in metadata remediation.

**Ongoing Reorganization of Technical Services**

As reported last year, Technical Services has been expanded to include responsibility for collections overall. Chris Cronin’s portfolio includes the following divisions with a total of approximately 140 FTE: Avery Architectural and Fine Arts Library, Rare Book and Manuscript Library, C.V. Starr East Asian Library, Original and Special Materials Cataloging, Collections Acquisition and Description, Collection Development and E-resources Management, and the Libraries’ Assessment program. This broader structure has already begun to facilitate work and information sharing resulting in greater collaboration and fewer unnecessary delays and hand-offs. The year saw retirements, resignations and new hires. New hires of special note include Kevin Schlottmann as the new Head of Archives Processing and lead on ArchivesSpace implementation, and Timothy Ryan Mendenhall as Metadata Librarian replacing Amber Billey. OSMC is in the final stages of hiring a new Special Collections Cataloging Librarian. Avery Library hired Lena Newman as Special Collections Librarian.

**Monograph Acquisitions & Cataloging**

As part of the Libraries Strategic Directions Initiatives, Technical Services workflows are being flowcharted and examined with an eye on increasing efficiencies. The impact of the "Discovery to Delivery" (D2D) Shared Collection at ReCAP has started to influence our handling of duplicate materials and in the near term these considerations will be reflected in our approval profiles as they move to more collaborative models. We have also been training staff and selectors to search the shared collections to avoid duplication. Cataloging has undertaken several cleanup projects as copy matching algorithms have surfaced problems in our records such as language or country of publication codes.

An increasing number of our approvals are arriving with MARC records. Since last year we now receive Albanian, Bosnian and Romanian materials with MARC records (Botimpex, University Press and Derex). We also now receive MARC records for Russian, Armenian and Croatian materials after having switched vendors to NKBooks and EastView for those languages.

Columbia's participation with the Ivy Plus Technical Services Group has continued. The group has been involved with comparing use of OCLC services and discussing formatting retention notes for print collections. It is now beginning to consider Metadata standards for Web Collections and will also be providing input on specifications for a shared index for Borrow Direct libraries.

**Program for Cooperative Cataloging**

Cataloging staff are actively engaged in PCC committees and task forces that are developing the infrastructure for the transition from MARC to BIBFRAME. Kate Harcourt is on the PCC Policy Committee, and worked on a small group to draft the 2018-21 Strategic Directions, Melanie Wacker
serves on the Standing Committee for Training. The PCC ISNI Pilot and is an advisor on the URIs in MARC Task Group, and Robert Rendall represents CONSER for both Columbia and Cornell.

**ArtFrame**

Melanie Wacker, Robert Rendall and Ryan Mendenhall have continued their participation in ArtFrame which is part of the LD4P Project. The group completed ontology modeling in collaboration with the Rare Materials Group lead by Cornell. This resulted in a combined Art and Rare Materials BIBFRAME Ontology Extension (arm). Other outputs include art-based use cases and a number of recommendation documents outlining decisions for specific modeling cases. Currently, the group is working on developing an application profile to support template development in VitroLib.

**Cataloging and Metadata**

Catalogers are involved in new initiatives coming out of Strategic Directions. Format specialists are doing a second inventory of hidden collections to measure progress and test reporting metrics. Susan Summer is leading a project to catalog unprocessed Middle Eastern books using selectors to find copy and provide metadata for original cataloging when needed. This strategy will be applied to other areas where records are unavailable or inadequate and cataloging departments lack language skills.

This year we expect to concentrate on workflow efficiencies. Progress in this area will provide staff for work on Hidden Collections. Small changes in DVD housing for example freed up staff time as well as achieving noticeable savings in the materials processing budget.

**Re-evaluating the Local Print and Electronic Collections**

In the coming year we will be re-evaluating the remaining print journal collections for possible replacement by electronic format when suitable online content exists. We will also be exploring opportunities to experiment with “Just-in-Time” approaches to collection development, particularly around DDA for eBooks and streaming video.