2CUL Technical Services in 2015
(A Work in Progress)

INTRODUCTION:

Any projections as to the shape of technical services for 2CUL in 2015 will necessarily be speculative and imprecise. The future will be determined as much by developments in publishing and technology as by our decisions and actions. Some trends may be predicted with confidence, such as the continuing shift from print to digital resources, but the pace of change remains uncertain. In other cases, such as the methods that will prevail for purchasing e-books, even a general pattern is difficult to foresee.

Given these uncertainties, we describe a "best guess" scenario for where we'll be in five years, but also outline optimistic and pessimistic alternatives. In this case, optimism projects changes that will bring savings in current technical services functions, allowing redeployment to other activities. Pessimism predicts fewer savings and/or the need for additional resources.

To understand potential changes in technical services resources, we need to consider the types of content our libraries will be acquiring in five years, the methods by which that content will be purchased or licensed, and the ways metadata to manage those resources will be created and maintained. These areas are considered in the sections that follow.

SYSTEMS ENVIRONMENT:

2CUL will have moved to a new, shared library management system (LMS), either through a commercial vendor such as Ex Libris, or using open-source software such as Kuali OLE. (The term LMS is used as a convenient shorthand, though the system will differ in significant ways from the current LMS.) Descriptive metadata will be stored in a single, shared database (Knowledge Base, or KB) – any resource will be "cataloged" once, for Columbia, Cornell, and potentially other users of the system. The Knowledge Base will contain additional metadata that is not library-specific: most vendor information, contents of electronic packages, linking syntax, connections for federated search, etc.

The Knowledge Base will be localized with information as to specific library holdings (location, call number, and barcode for physical items; license and access terms for e-resources; invoice and payment information; etc.) This information may be hosted or reside locally.

These data structures will allow the libraries within 2CUL to function separately or as a unit for groups of resources as appropriate. At a high level, Cornell and Columbia will act as "locations," with individual libraries serving as sublocations. Resources that are jointly purchased or licensed will be managed through a super-location, such as 2CUL or NERL. Staff accounts will allow access to information pertinent to either library or both, depending on the functions performed.

Optimistic alternative: An effective, affordable, well-maintained new system is available within three years. The Knowledge Base supporting e-resources is well maintained by a central agency (such as OCLC, SerialsSolutions, or ExLibris) at reasonable cost, supported by improved data exchanges with publishers and vendors, and greater adherence to standards.

Pessimistic alternative: Commercial systems fail to deliver needed functionality at affordable costs, and open source alternatives require significant local investment to develop and maintain. Knowledge Base maintenance requires continuing and substantial involvement from library staff to supply missing data and correct errors. Libraries continue to rely on pieced-together systems and data from multiple sources.

CONTENT, ACQUISITIONS, AND ACCESS
Domestic Market for Commercial Monographs
Content (~70% ebooks; ~30% print)
Processing (~85% by third parties; ~15% locally)

In addition to the degree to which the libraries will embrace what is possible from technical and market perspectives, there are two main issues that will impact the staffing resource requirements in the future technical services operation—a.) what percentage of new acquisitions will be electronic rather than tangible and b.) what percentage will be processed by a third party rather than by technical services staff at the 2CUL libraries. The scenarios presented below reflect what we believe to be the optimistic outcomes that could be achieved if technology continues to evolve at its current pace, the markets exist for a digital-based strategy, and adoption both by the user communities and the library decision-makers support an aggressive shift to a digital-preferred collection model.

<u>E-books (~70% of all monographs acquired)</u>
The dominant means to acquire over half of the content newly published in the domestic market will be blanket/subscription plans for ebooks. Selectors will create subject/publisher profiles (how fine-tuned the profiles will be will depend on the pricing model—it could be more attractive from a cost/benefit perspective to acquire a large package than to fine-tune a custom package). The vendor/publisher/aggregator will turn on access to ebooks that match the profile within the KB. Payments will be made ahead of the publishing cycle, just as we do now for journal subscriptions, but will not be tied to specific books (published units). Descriptive metadata for the content will be updated in the appropriate venue (see Descriptive Metadata below). For several publishers, print copies for titles included in the acquired e-packages will be available for a relatively low unit cost and printed on demand, but for the most part, it will be the user who buys his/her own copy taking advantage of the institution's discount. The Library will be focused on building a collection of/guaranteeing access to commercially produced econtent, not print. To a lesser degree (30-40% of ebook acquisitions), some selection of ebooks will be patron-driven based on their use of titles, the records for which we will pre-load into our discovery databases.

Print monographs (~30% of all monographs acquired)
The remaining acquisitions will consist of print monographs not already owned (and not available in a digital format) and will continue to be accessed, one-by-one, but mostly driven by faculty or other community user recommendations—in other words, a patron-driven print acquisitions plan. Our users will discover such titles as part of their searching of a scoped OPAC that we will have prepopulated with metadata for titles we can quickly acquire (within three days). The purchase requests will be forwarded to the correct processing stream through the 2CUL tool, POOF! to expedite acquisition.

Most of the new tangible materials that are purchased by the library will be shipped ready for shelving, including any related descriptive metadata. Printing services such as Ingram's Lightening Source will be able to print titles shelf-ready when producing tangible books from digital files. Other print materials will be processed by the vendor supplying the books. Even so, a small percentage (~5-10%) of non-digital items will need local processing to support shelving and security.

Remaining Market for Monographs
The dominant means for acquiring new monographs from beyond the US and Canada will be through approval/blanket plans (55%) with the remainder acquired through title-by-title ordering .

Commercial publications from the **European market**, while continuing in a mostly print format (~80% acquired as print), will be processed mostly by vendors. We expect to need local processing for about 25% of all European titles acquired, regardless of format.

For the **East Asian market**, we anticipate up to 50% of new receipts will be published in electronic format. We expect over 90% of all East Asian titles could be processed by a third party.

The **Latin American market** will be about 10% electronic. About 75% of all Latin American titles could be shipped already processed, leaving 25% to be completed in-house.

For the **rest of the world**, we imagine about 10% of the monographs to exist as ebooks, but expect most of the processing to continue to be done in-house.

Optimistic alternative: in the next five years, we can't imagine a more optimistic scenario than described above.

Pessimistic alternative: while there will be more content available in electronic format than today, either the market or the institution might not embrace the optimistic level of saturation projected above. Even so, the infrastructure to support both a robust digital and print collection will need to be supported.

Market for Serials
By 2015, we expect access to the world's serial literature (journals and periodicals) to be 80% electronic. The "electronic + print" option will largely disappear, except for a few

titles with an embargo on recent online issues. The remaining serials will continue to be acquired in print and local holdings for titles not available electronically will be recorded in our inventory systems.

Licensed Resources
The licensing process will have become much more generic and will be handled by a third party, regardless of packages purchased at the individual institutions. Such a third party might be a consortia of two or more institutions, though will tend towards the larger model currently embodied by NERL and Lyrasis. Anything not supported by a NERL-like hosting service will be handled by a single 2CUL Electronic Resources service. Beyond 2CUL, we will continue to pay an annual fee (though it will be somewhat larger than at present time) to cover the staffing of our hosted licensing workforce. Licensing will encompass all forms of media, not differentiating between e-books, e-journals, or e-streams of audio or video.

Authentication and rights management will occur in the knowledge base (KB) mentioned above under the Systems Environment. Rather than maintaining a separate Electronic Resource Management system (ERM), the activity and data will be aggregated in a common KB, and in most cases, maintained by the third parties from whom we make purchases of electronic content.

Optimistic alternative: Licensing at the consortial level will cover license terms, price, and content for a high percentage of electronic resources, significantly reducing the need for local management. Most remaining e-content will be managed through joint licensing for 2CUL.

Pessimistic alternative: Economic pressures combined with differences in academic programs will limit the potential for consortial licensing. License terms will be more standardized than at present, but new business models will continue to be offered and adopted at different rates by libraries, requiring continuing investement of resources for negotiation and maintenance.

Access to Content
Academic users (undergraduates, graduates, and faculty) will own or have ready access to affordable e-reader devices that fully support the functionality sought by most users—easy ingest, maneuverability, note taking, ubiquitous access, and ADA compliance. The current market of readers, the supporting platforms, and the storage and display standards will have matured to the point that users will be fully satisfied with an affordable electronic device-supported access experience.

As mentioned above, many of the agreements for access to electronic content the Library enters into with information providers will allow users to purchase personal print copies at a fraction of the list price.

ILL/Document Delivery
The mass digitization projects by such entities as Google, the Internet Archive, and publishers will have eliminated the need for much of the delivery of tangible items between

institutions. For items in copyright, it will be more economical to 'rent' content from a commercial entity than to retrieve, ship, receive, and reshelve tangible items. Local users will be offered a choice of free access to locally licensed or public domain electronic content or a fee-based service to request tangible copies of content available locally electronically. Libraries will continue to provide free access to content that either does not exist in electronic form or is not available locally.

DESCRIPTIVE METADATA:

Commercial publications:

Two groups of resources will operate in very different metadata environments.

Textual digital resources will be metadata rich: readily discoverable by full-text searching and text-mining for metadata generation, leveraging metadata created to support the supply chain. Commercial audio and visual digital resources will not benefit from the advantages of full-text searching, but metadata for discovery will nevertheless be available through the supply chain.

Analog resources, by contrast, will be underprivileged with regard to metadata. Most will have descriptive information available on the web, but not in sufficient depth to compete with digital resources for discovery and attention.

For most print books, basic descriptive information will be supplied by the publisher and/or vendor. Many libraries, including the Library of Congress, will continue to catalog these books using standard rules and vocabularies (AACR, RDA, LCSH, etc.) Research libraries, including 2CUL, will continue to accept these records "as is."

For those works lacking copy, the current cataloging model will be replaced by a new one. Catalogers will continue to create basic descriptive information where none exists, but will largely accept such information when supplied, without adjustment. They will augment the description with information available on the web, with particular attention to contents and summaries. For authors and other named entities, they will create and/or augment information in linked data repositories. Subject vocabularies will be applied based on the expected audience, expressed as linked data to take advantage of linkages between ontologies. Subject analysis will be deeper and fuller than at present, to compete with the deeper access available for full text resources.

The resulting metadata environment will be diverse and inconsistent, in a state of transition. For new books, about 50% will have very rich access based on full text, only loosely related to library cataloging practice. Another 30-40% will continue to be described (by LC and other libraries) according to current library standards. Some 10-20% will make use of new modes of cataloging, designed to enhance discovery for specific user communities, in parallel with the types of access available for full-text resources.

These changes will allow some savings in staff resources, potentially offset by greater fees for metadata that is externally supplied and maintained. Savings on metadata for most US/UK publications will be minimal, since costs in this area are already very low. However, access to these materials will be greatly enhanced by full-text searching. Staff costs for processing foreign materials will be reduced somewhat thanks to greater availability of externally supplied metadata. Unit costs for original cataloging (as applied within 2CUL) will be reduced slightly.

Optimistic alternative: A more rapid shift to electronic books outside the US and Europe could bring further savings, as metadata that is now manually created is replaced by full-text searching. For analog publications, further savings could be achieved by reducing efforts to supply library-created metadata, relying on external sources and web searching for discovery, and linked data through WorldCat for access via the library.

Pessimistic alternative: Pressure to recover the costs of metadata creation could increase the price of externally-created metadata, offsetting staff savings. Community pressure to adhere closely to library cataloging standards could reduce potential savings (and benefits) in original cataloging.

Metadata Services:

Demand for high-quality, high-performing metadata for non-commercial publications will increase as special collections are digitized, institutional repositories grow, and new forms of born-digital resources emerge. Metadata will be based on a diverse set of community standards appropriate to the content of the resource and primary audience, relying on linked-data schema for cross-sector discovery. Some metadata will be generated automatically from the resources, some supplied by authors, and some created by part-time subject specialists (e.g., graduate students.)

Catalogers will play a role in metadata design for specific projects and collections, and in reviewing metadata for potential new entries in linked data repositories (e.g., name authority files, subject vocabularies.) For selected resources, catalogers may do original cataloging, particularly where integration with commercially-published resources is important.

Resource allocation needs for metadata services are highly speculative and to some degree discretionary. The emphasis for 2CUL is likely to be on retaining a cadre of metadata specialists/catalogers with a range of language and subject expertise, and employing those skills to best effect, rather than on expanding or reducing staff.

Optimistic alternative: Continuing improvements in web search and discovery may reduce the need for standardized metadata. Contributions to metadata by scholarly users of digital resources may reduce the need for initial rich metadata creation.

Pessimistic alternative: Investments in digitization and creation of unique digital resources (within 2CUL and beyond) may outpace investments in metadata creation,

increasing pressure for large research libraries to invest in metadata enrichment beyond their own collections.

IMPACT ON 2CUL:

Probable effects of these changes on technical services in 2CUL can be depicted in broad strokes, but are difficult to quantify. Broadly speaking, we anticipate:

- Modest reduction in the staffing devoted to managing print resources (Monographic Acquisitions, Copy Cataloging, and Serials Acquisitions.)
- Modest expansion in the staffing devoted to electronic resource management (Electronic Resources.)
- Merger of the staffing devoted to Metadata Services and Original Cataloging, with little change in total staffing for these functions, but considerable change in the work performed.

Deep collaboration within 2CUL is likely to focus on three areas:

1) Joint management of electronic resources; i.e., a single administrative unit managing ordering, licensing, payment, renewals, data maintenance, and problem resolution for e-resources acquired jointly or separately.
2) Shared use of expertise within original cataloging and metadata services; i.e., applying language, subject, and domain expertise (such as deep knowledge of specific formats and standards) to relevant content regardless of ownership.
3) Co-development and application of efficiencies in metadata management, including tool development (such as the current work on POOF), data loading, metadata transformation and enrichment.

Within technical services, 2CUL also has the potential to play a leadership role in generating efficiencies for the broader research library community through such actions as: joint approach to foreign vendors to propose new approval plans and cataloging services, joint exploration and adoption of new modes of cataloging, etc.

Scott Wicks (Cornell University Library) scott.wicks@cornell.edu
Robert Wolven (Columbia University Libraries) wolven@columbia.edu
September 15, 2010