Reconsidering silos

Received (in revised form): 24th May, 2022

Tammy Troup
Digital Preservation Librarian, Iowa State University Library, USA

Tammy Troup is a digital preservation librarian at the Iowa State University Library, where she is responsible for development of policy, oversight of technical procedures and assessment of bitstreams in support of the long-term retention of digital objects. She holds a master's degree in library and information science from Kent State University, where she studied digital preservation.

Iowa State University Library, 701 Morrill Rd, Ames, IA 50011, USA
Tel: +515-294-3669; E-mail: ttroup@iastate.edu

Heather M. Campbell
Metadata Management Librarian, Iowa State University Library, USA

Heather M. Campbell is a metadata management librarian at the Iowa State University Library, where she oversees the creation, normalisation and quality assessment of metadata for the library's digital collections. She holds a master of science in library and information science from the University of Illinois at Urbana-Champaign.

Iowa State University Library, 701 Morrill Rd, Ames, IA 50011, USA
Tel: +515-294-2157; E-mail: campbe11@iastate.edu

Chris Dieckman
Cataloguing and Metadata Librarian, Iowa State University Library, USA

Chris Dieckman is a cataloguing and metadata librarian at the Iowa State University Library, where he is responsible for overseeing the cataloguing of library materials and assisting with metadata creation and enhancement. He holds a master's degree in library and information science from Kent State University, with a specialisation in cataloguing and metadata.

Iowa State University Library, 701 Morrill Rd, Ames, IA 50011, USA
Tel: +515-294-9632; E-mail: dieckman@iastate.edu

Laura Sullivan
Digital Collections Librarian, Iowa State University Library, USA

Laura Sullivan is a digital collections librarian at the Iowa State University Library, where she manages the digital collections platforms and digital surrogate files. She holds a master of arts degree in library and information studies from the University of Wisconsin–Madison, with a specialisation in archives and records administration.

Iowa State University Library, 701 Morrill Rd, Ames, IA 50011, USA
Tel: +515-294-9680; E-mail: lsulli12@iastate.edu

Abstract  While organisational silos are often viewed as an obstacle to overcome, this paper argues that they are a necessary feature of the workplace environment and should instead be approached in a manner that maximises their potential. This case study describes an approach used by three specialised areas at an academic library — cataloguing and metadata services, digital collections and digital preservation — to retrospectively manage the technical debt incurred during the COVID-19 pandemic.
In response to the sudden shift to remote work, library workers developed an action plan to continue an in-process migration with minimal disruption; however, this necessary action significantly exacerbated an already complex file storage system. Although retrospective management of technical debt can be complicated by the management needs of multiple stakeholders, staff members in several areas of specialisation approached the project by identifying needs related to the subcategories of a digital object entity — the intellectual entity, representation entity, file entity and bitstream entity — and developed high-level solutions to meet these needs. This approach exposed three types of silos (context, schemas and processes) and related communication challenges, which can complicate cross-team collaboration. Yet by reconsidering organisational silos as interconnected units of specialisation, staff members successfully applied specialised knowledge, advocated for their management needs and collaborated to resolve the technical debt.

KEYWORDS: professional specialisation, communication in organisations, teams in the workplace, academic libraries, digital preservation, organisational effectiveness

INTRODUCTION

Modern information systems are complex environments that ‘need to create some structure to handle this complexity’.\(^1\) In academic libraries in particular, complexity results in specialised work silos to meet organisational objectives and extend roles and responsibilities that build librarianship portfolios ‘distinct from the scholarship and service that is required for rank promotion’.\(^2\) The case study described in this paper shows that organisational silos, rather than acting as barriers to collaboration, can be useful and efficient.

The authors are librarians representing different departments — digital preservation, digital collections and metadata services — within a large academic library at a research university. In 2021 they formed an ad hoc team to package digital objects and transfer those packages to fixity-monitored storage. This effort, called the digital surrogate bagging project, was initiated and led by the digital preservation librarian — a staff member without positional authority to direct and manage the other team members. To complete this project successfully, the authors relied on the ideas outlined in the Reference Model for an Open Archival Information System (OAIS) and the PREMIS data dictionary.\(^3^,\(^4\) Despite belonging to the organisational silo of digital preservation, the OAIS model and vocabularies and the data models of PREMIS data dictionary framed objectives that moved the project forward.

As the team used these digital preservation models to collaborate on the project, they reconsidered the negative connotation of silos. The resulting approaches to their work processes and communications maintained the agency and autonomy of the multiple silos in their respective areas. Reconsidering organisational silos as units of specialised knowledge, functioning separately while remaining interconnected, moves information organisations away from a metaphor of destruction and towards one of stewardship. This paper, a case study of the surrogate bagging project, posits that cross-functional teams can find benefits in reworking or reimagining silos, rather than destroying them.

LITERATURE REVIEW

Silos, when defined as isolated groupings or departments that function in a way that hinders cooperation, have a negative connotation.\(^5\) Within library and information
science literature, silos are often targeted as structures to be torn down. As Kowalski has argued, departmental or functional silos inhibit organisational goals, limit efficiency and creativity and hinder professional development.6 Similarly, Massis justifies the removal of silos to prevent confusion and the appearance of a ‘disjointed jumble of independent silos rather than an interdependent whole’, adding that a goal of the academic library should be to ‘create cohesiveness among the staff at every level’.7

Silo destruction is an attempt to solve multiple problems related to project management, workflows, communication and collaboration. Birrell and Strong describe a community-building approach to breaking down silos that includes cross-team training and knowledge-building as a project objective.8 Kowalski and Forsten-Astikainen et al. likewise advocate for breaking down silo walls through such means as shared procedures and workflows, open communication, communities of practice and mentorships and cross-training.9,10 MacDonald and van Duinkerken adopt a more emphatic metaphor, referencing the idea of ‘creative destruction’ — the replacement of traditional processes and organisational structures with new ones fostered by an entrepreneurial culture — as a means of survival.11

Libraries’ efforts to break down silos reflect trends in the business world. Calls to bust silos are common ‘both in practitioner and scholarly journals’.12 Business researchers have cited the negative effects of silos, including rigid organisational boundaries, a lack of cohesion, internal competition and underperformance.13,14 To counteract these harmful trends, leadership consultant Henman advises leaders to ‘remove their silo-building behaviours and replace them with silo-busting decisions’.15 However, not all business literature advocates silo destruction — pragmatists recognise that silos exist and will continue to exist. Business leaders Vantrappen and Wirz, for example, suggest instead that silos can be beneficial, provided that organisations approach them effectively.16 Specifically, the authors argue that to minimise negative side effects, organisations must build bridges (e.g., through shared values and networking) and institute checks and balances (e.g., through governance and intervention mechanisms).

Indeed, a variety of thinkers have advanced the view that silos are not only inevitable but a necessary structure in an increasingly complex environment. Legal scholar Benkler’s exploration of the networked information economy highlights the importance of modularity in the information production process, which ‘must effectively integrate widely dispersed contributions, from many individual human beings and machines’.17 In plainer terms, Tett, an anthropologist, writes that ‘the modern world needs silos, at least if you interpret the word to mean specialist departments, teams and places’.18 Vantrappen and Wirz agree, stating that silos aggregate expertise, assign accountability and provide identity.19 Even Henman, an advocate of silo-busting, recognises the need for specialisation: ‘Traditionally, people have focused on a division of labour in work groups; however, top-performing teams require a shift in paradigms — a movement toward and focus on a division of knowledge’.20

Although some library leaders have continued to advocate for the destruction of silos, library workers at some organisations have embraced a considered approach to silos to focus on the exchange of knowledge across areas of specialisation. Barlow and Bocko have discussed their experience of building a culture of knowledge exchange to extend their roles and support collaboration.21 Similarly, in addition to developing a culture of knowledge exchange across silos, Alagna et al. have offered guidance on developing realistic workflows, including those that
cut across multiple units. Examples of cross-silo workflows include the structured workflows developed by Oberbichler et al., the Linear Reciprocity Model designed by Pryse and the project management approach of Bertoldi et al. The Linear Reciprocity Model, for example, relies upon a ‘simplicity of process within each module [to secure a] barrier-free flow of information’. The Bertoldi et al. team’s focus on human infrastructure led them to recognise the benefit of ‘the shared expertise of its members, especially when dealing with diverse collection items and differing organisational structures’.

Working across established silos requires staff to build lines of communication and structure for knowledge sharing. However, as Kris Joseph has asserted, ‘the arbitrary creation of digital silos is a move in the wrong direction; what remains is to explain how organisational structures — whether they are digitally-centred or not — can affect the creation and distribution of knowledge within the library’.

Reconsidering silos as units of specialised knowledge, interconnected through collaboration, may help to foster more efficient and inclusive knowledge sharing within an organisation. While retaining the destructive metaphor of ‘breaking down walls’, McGovern proposes radical collaboration — a form of collaboration that requires communal commitment and interactive engagement, as a way to build ‘sustainable, inclusive communities that are able to solve problems together by leveraging cumulative strengths’.

**SILOS**

The organisational silos addressed in this paper include the traditional departments of Metadata Services and Preservation Services and the newly developed Digital Scholarship and Initiatives. Professionals in these organisational units possess familiarity with various technical tools and schemas, as well as specialised experience based on past projects. Although each unit has a high-level understanding of the others’ work, they all rely on their colleagues to clarify and apply in-depth knowledge particular to their areas — the disciplinary knowledge that, as O’Neill and Stapleton observed, ‘is sticky, embedded and situated’.

Despite MacDonald and vanDuinkerken’s warning that ‘creative destruction’ is ‘essential to the process of developing an entrepreneurial culture’, the cross-functional team did not find it necessary to abandon traditional silos. Indeed, the traditional organisational silos provided team members with a broader network of specialists, and they extended the capacity of individual members. Thus, by reconsidering silos as units of specialised knowledge, interconnected through the sharing of information as needed, staff gained an efficient model for collaboration on complex tasks.

Nevertheless, in the course of completing the surrogate management project, the team navigated knowledge gaps across their areas of specialisation. These knowledge gaps did not result from the siloing of expertise within the organisational units. Rather, the expertise required to fill these gaps crossed organisational boundaries. The specialists thus characterised the expertise needed in these cross-functional areas as additional types of knowledge silos — ones that did not map to organisational areas. They named these knowledge silos context, schema and process silos.

**Context silos**

Within an OAIS model, digital object management requires identifying appropriate context for a designated community. System maintainers must either provide a minimal amount of context for a specific community or purpose, or they must offer more context for broader understanding. System maintainers thus navigate multiple
layers of contextual information and make decisions on its capture and retention based on community needs and expectations. This is not straightforward work, as information is ‘inherently relational; it is always context of, about, or surrounding something’.33

Moreover, determining context can be fraught with tension as context may be meaningful to one designated community and worthless to another. Context to understand content is typically prioritised and carried forward as metadata, while context about business practices is not prioritised. Although past practices can be discerned (e.g., by reviewing data structures, documented workflows, training guides or project descriptions), contextual information about digital surrogate management is not always captured systematically.

**Schema silos**

Related to the idea of context is the capture of data, including metadata and the management of metadata as an information resource. Digital objects may be accompanied by metadata standardised for consistent capture of administrative, descriptive, technical and preservation information. With regard to digital preservation, the use of metadata is a crucial component of an archival information package, as it ‘facilitates and documents the storage and transfer of the bag’.34

Metadata schemas provide guidelines for the structure and presentation of metadata, which can be expressed as plain text, in tabular format or in programming or markup languages, such as XML. Schemas are facilitating documents, developed and maintained by communities of practice, that ‘[define] syntactic, structural and semantic features of metadata used for a specific application’.35 Metadata support contextual understanding, although not every context may be supported by a given schema. In general, the development of custom schemas is discouraged; indeed, ‘one of the challenges of a [metadata] manager is to weigh up the relative benefits of developing a schema that suits a particular application … and the use of a commonly accepted standard that allows for interchange of data’.36

Identifying and documenting core fields, controlled vocabularies and input standards, as well as assessing interoperability between standards, are highly specialised activities, which require familiarity with various authorities, community resources, local guidelines and encoding or markup practices. However, decisions regarding the implementation of various metadata standards and practices change over time and are not always well documented.

**Process silos**

The use of digital technology is an inherent feature of digital object management and system maintainers develop varying levels of familiarity with the tools used for maintenance. Despite the ability to switch platforms and intuit new technologies, system maintainers can develop silos around the use of tools. Familiarity with technology can be relied on to quickly overcome learning curves, expedite workflows, hand off tedious tasks, and even prevent digital fatigue. The use of familiar tools provides maintainers with autonomy over their processes and staves off exhaustion that is introduced with ‘new organisational systems and the mandatory training, adaptation and use’.37

In addition to silos developing around technology, as ‘generations of practice follow generations of technology’, past practice and use of obsolete technologies become intrinsic to a given knowledge silo.38 This knowledge can be as simple as understanding changes in file extension protocols or recognising that a record was created prior to addition of elements in a given schema or development of a content standard. Sharing this knowledge across organisational silos
is an example of the radical collaboration recommended by McGovern.39

Ideally, process silos are a means of extending the capacity of an organisation: they support efficient assessment and validation and the use of ‘standardised processes remove barrier formation, duplication, data entropy and redundancy’.40 If approached correctly, process silos follow ‘good digital practice’, which is ‘cumulative, iterative, responsive to organisational and technological change, inclusive and open’.41 It is important to acknowledge that process silos can become barriers. Kowalski calls for silo walls to be broken down when ‘an employee or department considers a task “not my job”’.42

Cross-silo communication
In addition to the existence of context, metadata and process silos, the authors navigated communication challenges. The organisation is in the early stages of self-awareness as an archival information system with digital preservation as a goal. A system-wide approach to digital preservation offers staff an excellent opportunity to develop cross-silo methods of communication and new frameworks for identifying system needs. While systems thinking is both an organisational method and a philosophy with a distinct vocabulary and design process, systems maps may be overwhelming without an introduction to systems thinking.43

As some staff members are becoming familiar with the long-term management of digital resources, they balance a growing need to understand the ‘big picture’ — the rationale for digital preservation — with the everyday application of their specialisation to meet near-term objectives. This incomplete understanding does not always lead to the best solutions.

Another potential area for miscommunications between team members is jargon or language specialisation, particularly with regard to metadata and the technical systems used to create and manage metadata. Shifting activities from creating metadata to managing metadata requires ‘a detailed understanding of the functions of metadata in relation to particular user contexts, just as the creation of metadata does’.44 The more emergent area of digital preservation can also be rife with buzzwords and overly detailed explanations. Del Pozo et al. notes that ‘the field of digital preservation is like a monastic order, whose chanting is gibberish to those outside of its walls’.45

Finally, the use of online meetings, e-mails, project management tools and deskside chats results in both synchronous and asynchronous communication. Synchronous communication is ‘the interaction in which participants can communicate in real-time — without significant delays’, while asynchronous communication is ‘a communication exchange that does not take place at the same time for its participants’.46 If team members have differing expectations regarding communication tools and styles, inconsistent communication on project objectives can occur, leading to flawed assumptions or unmet expectations.

With full awareness that knowledge silos and communication challenges would add a layer of complexity to the project, the team applied a strategy for organising the division of labour, which relied on a model familiar to digital preservationists. This case study describes the use of a digital object management framework to identify each entity of a digital object and to identify methods of management for long-term access.

CASE STUDY
Background
The university library began digitising resources from the general collection and special collections as early as 2000. These digital objects were presented to the public on Luna Imaging and later migrated to CONTENTdm. Unstructured files were also
managed on centrally managed servers and network attached storage devices.

Decades of object creation and management strategies resulted in drives with duplicate files, prioritised TIFFs managed alongside access JPEGs, metadata structured for single-use ingest, deep directories that mirrored physical file structures and complex file-naming conventions.

In 2019, the university library planned a migration from CONTENTdm to Islandora, and staff copied, edited and organised digital objects in support of this migration. In March 2020, the university pivoted to remote work in response to the COVID-19 global pandemic, and staff copied all digital files to a network location with a remote connection for off-site access. These actions, while necessary, complicated the already complex file storage system. The resulting work needed to manage these files is technical debt — the consequences of deferring long-term priorities in order to enable short-term gains.47

To resolve this technical debt, after migration, staff set a goal of implementing a strategic approach for the management of preservation files. The following objectives met this goal: identify digital files that mirrored files uploaded to Islandora, harvest administrative and descriptive metadata, extract technical metadata, develop preservation metadata and package and transfer files to secure, stable, fixity-monitored storage. This work required the support and input of multiple staff members in several areas.

The project manager — the digital preservation librarian — drew on their area of expertise to consider the specialisations needed for the project. Digital preservation is informed by the OAIS model and vocabulary and the PREMIS Data Dictionary, which includes a schema for the management of digital resources and a model for management.48,49 This data model identifies four distinct entities required for digital object management: objects, rights, agents and events. The object entity in particular is expanded with subcategories: intellectual entities, representations, files and bitstreams. These categories became a useful guide for thinking about the stewardship requirements of digital objects, and they guided the investigation and solutions development as well as the technical stack planning for the digital objects.

Investigation and technical stack plans

The order of investigation and identification of solutions proceeded from securing the bitstreams to discovering the bitstreams. The iterative query process (see appendix) required staff to communicate across their organisational silos. This process exposed system needs (hardware, software, workflows, metadata, etc) as well as past operational practices and documentation gaps. The planning process proceeded across the layers of the digital object as follows:

- **bitstreams**: securing bitstreams;
- **digital files**: identifying digital files;
- **representation entity**: identifying item-level and relationship information resources; and
- **intellectual entity**: identifying a collection-level information resource to support inventory.

In Figure 1, the high-level tasks required for each phase of inquiry are mapped to the four subcategories in the digital object data model, represented as columns. The areas of specialisation of each staff member, represented as rows, intersect with one or more of these subcategories. Arrows connecting the high-level tasks show multiple interconnections across subcategories, while dotted lines between the areas of specialisation represent the porosity of functional divisions. By contrast, Table 1 identifies the deeply specialised knowledge required for each high-level subcategory. This level of specialisation is not apparent in the high-level diagram, as practitioners typically
share these details only when needed for clarification.

Figure 2 illustrates the technical solutions identified during the investigation process. In this figure, the subcategories of the data model are represented as columns with distinct technical solutions identified by the areas of specialisation. Collection-level administrative metadata are captured with MARC records in Alma. Staff harvested item-level administrative, descriptive and structural data from Islandora, extracted technical metadata with FITS and structured preservation data in a CSV with PREMIS.

**Context silos**

Each area of specialisation identified and provided specific contextual information about the digital surrogates. The digital collections librarian had the most familiarity with the collections, and focused primarily on the surrogate files themselves — providing historical and background context as needed and delivering decisions for the files of record. Prior practices that impacted this case study included a practice of managing active production files, file restructuring for a planned migration and an emergency response for remote access; contextual information was often undocumented or with an unclear documentation location. The digital collections librarian provided a second voice for decisions made within the other areas of specialisation, and they shared contextual information during the preservation file isolation process. Input was given most often when compromises needed to be made, such as including files when they constituted less than 2 per cent of the total collection or leaving files in production if the percentage was greater.

**Silos and solutions**

While the project progressed, the participating specialists confronted gaps in understanding caused by the three types of knowledge silos — context, schema and process — discussed above. As plans and investigation unfolded, they identified solutions for bridging these gaps. While maintaining their areas of specialisation, they adopted communication strategies that built interconnections among the silos.
<table>
<thead>
<tr>
<th>Area of Specialisation</th>
<th>Content</th>
<th>Context Silo</th>
<th>Specialised Processes</th>
<th>Specialised Tools</th>
<th>Process Silo</th>
<th>Schema Silo</th>
</tr>
</thead>
<tbody>
<tr>
<td>General context</td>
<td>DACS, LCSH, RDA</td>
<td>Box, Microsoft Office, File Format characteristics, Adobe Acrobat</td>
<td>Recording MARC metadata, Batch converting CSV metadata, to MARC Metadata display</td>
<td>Islandora, OpenRefine, Python</td>
<td>Pomodoro, Crosswalk</td>
<td>MODS, Schema.org, ISO 8601, MODS, RELS-INT, RELS-EXT</td>
</tr>
<tr>
<td>Specialised context</td>
<td>Alma, MARC, CSV, Import</td>
<td>Alma, Crosswalk</td>
<td>Quality assessment and control methods, Research practices, Structured data editing, Python programming</td>
<td>Islandora, Crosswalk</td>
<td>Targeting, Crosswalk</td>
<td>MODS, BagIt, Tags</td>
</tr>
<tr>
<td>Digital asset</td>
<td>Digital content and metadata, DACS, LCSH, RDA</td>
<td>DAMS, Google Drive</td>
<td>Quality assessment and control methods, Research practices, Structured data editing, Python programming</td>
<td>Islandora, Crosswalk</td>
<td>Targeting, Crosswalk</td>
<td>MODS, BagIt, Tags</td>
</tr>
</tbody>
</table>

Table 1: Knowledge silos.
Digital preservationists typically capture and carry forward files with a consistent record of integrity and reliable creation dates. As the data could not be reliably identified, preservation staff relied on the following data points to provide context and serve as significant properties — standard type-based significant properties, uniqueness, relationship with metadata (filename), collection completeness and format.

With benchmarks in place for identifying preservation files, preservation staff used the recently developed packaging guidelines to set final package expectations for bagged surrogates. As shown in Figure 3, bags conformed to a directory name that supported quick access to collection name, year bagged, platform and bag sequence and they contained consistent metadata files and content files. Staff removed working directories prior to bagging.

Metadata staff focused primarily on providing contextual information for the digital surrogates. Although after reviewing the expected bag contents, the metadata librarian advocated for inclusion of a high-level descriptive metadata file to accompany the bagged files, this high-level collection record provided helpful context during the implementation phase, and it kept preservation staff tethered to the content and the culture of stewardship of libraries and archives. In the future, this context can be added to the digital collection platform to describe the collections or used by future data managers.

Preservation staff carried forward management context using BagIt reserved metadata elements. These tags documented known administrative history including transfer procedures, storage environments, transfer agent name and location of related files (metadata packages).

**Schema silos**

The process of bagging files required the capture of technical and structural metadata to ensure authenticity and accurate rendering in the future, as well as administrative and descriptive metadata to facilitate future discovery and understanding. Although each librarian had experience working with metadata schemas, they had different levels of
familiarity with the schemas and mappings required for this project, including MODS, RELS-EXT, RELS-INT, PREMIS, MIX and MARC21.

Team members successfully identified methods for extracting various types of metadata from the surrogate files as well as from multiple datastreams in the library’s digital collections platform, Islandora 7.x.50 However, the library did not have a precedent for the capture of metadata to describe a collection (an aggregate of digital objects).

The team therefore had to select core elements that provided adequate contextual information about the digital surrogate collections. In a collaborative effort, team members settled on the Metadata Object Description Schema (MODS), a bibliographic element set already in use as a metadata schema for digital objects in the Islandora platform.51 The team selected seven descriptive and administrative MODS elements (title, originInfo/dateCaptured, abstract, identifier, language, note and type Of Resource) and determined content standards for populating the elements. Where metadata aligned to the selected elements could not be extracted or did not exist, metadata and digital collections librarians collaborated to create them.

Meanwhile, preservation staff captured preservation metadata — a type of administrative metadata created for a user community interested in the technical characteristics of the digital object, rights affiliated with the object, agent interaction with the object and events that have affected the object. The PREMIS data dictionary provides controlled terms and content standards, and specifies a preference for the automatic capture of such metadata.52 For the bagging project, preservation staff developed a manual-entry preservation metadata template with PREMIS elements to capture content for the following elements: objectIdentifier, preservationLevel, fixity, format, original name, contentLocation, relationship, relatedObjectIdentifier, relatedEventIdentifier, eventType, agentIdentifier, rightsBasis, copyrightStatus and copyrightStatusDeterminationDate. Staff shared the captured data with the metadata services team in order to populate MARC records.

After the capture and creation of the required metadata, the digital preservation librarian structured the data in preparation for said data to be added to the library’s local catalogue records. This instance of data handoff relied on preliminary work conducted by metadata staff and field mapping. Staff assessed and updated or created MARC records for the digital collections. Therefore, when adding the newly created preservation metadata, staff...
could concentrate solely on the germane 516 and 856 fields. A crosswalk was then developed using the Library of Congress’s documentation for crosswalking MODS to MARC21, tested on several existing MARC records and adjusted per the recommendations of various project stakeholders. Finally, the metadata and cataloguing librarian planned to use Ex Libris Alma’s CSV import feature to add the crosswalked preservation metadata in batch to the local catalogue records. However, most of the records had multiple 856 fields, and the batch-import feature does not allow for the addition of two fields with the same tag in a single CSV file. To work around this limitation, some 856 fields were assigned temporary tag numbers not used in MARC21, such as 857 and 858. After the metadata had been added to the MARC records, an Alma job was run to batch change these temporary tags to the correct 856 tag.

Process silos
Each area of specialisation used subprocesses to support other silos or to complete a task. During the planning phase, the digital collections librarian and the metadata staff needed to frontload a number of support documents used by the preservation staff. The digital collections librarian created a directory-to-collections map, generated a directory listing, and shared TIFF maps used for ingest. These documents supported the preservation staff’s process of preservation file isolation. In addition to preparing the high-level MODS records shared in an Excel document, metadata staff collected text files with persistent identifiers and prepared a Python script used by the preservation staff to harvest administrative, descriptive and structural metadata from the library’s digital asset management system.

Due to the considerable number of knowledge gaps regarding the preservation files, the planning phase of the project intentionally laid the groundwork for responsive and innovative process development during the implementation phase. The team provided the library’s digital preservation intern with an introduction to the planning process and package targets. The digital preservation librarian and intern worked collaboratively and independently to develop command line and open source software-supported procedures (see Table 1).

The intern developed expertise in micro-project management, process parallelisation, package assessment, file handling and bagging. Most importantly, they learned to rely on expertise within and across organisational silos and to communicate effectively and to gather contextual, schema or process-related knowledge.

Cross-silo communication
When developing a systems approach to digital preservation, documentation gathering and planning can look like busy work in an organisation focused on product delivery.

The digital collections librarian recognised that digital surrogate management for preservation would be relatively ‘low-hanging fruit’ and the team seized the opportunity to provide evidence of a minimally viable product while developing other system parts.

The digital preservation librarian prepared communication tools to support the project including a project brief, running meeting notes, basic package management frameworks, activity, files and metadata logs, preservation metadata templates and sample packages. The initial ‘project one-pager’ conformed to the project brief requirements used by the project planning division and included an outline of requirements to support an intern through the process of planning and implementing the project. The team quickly recognised that expecting an intern to support the planning process would be overwhelming and therefore expanded the plan with documented procedures.
The implementation phase involved weekly meetings and asynchronous chats between the project manager and the intern and the digital collections librarian and the intern. In general, staff found that knowing the basics of what was happening outside of their functional area was beneficial; however, the communication of most project details outside of the general goals was unnecessary for the successful completion of the project. Specialists provided information or developed a process when asked and were therefore able to concentrate a large portion of their time on other projects, prioritising the digital surrogate bagging project only when participation was necessary; specialists valued efficient communication and timely questions. As their knowledge developed, they saw how their contribution supported project goals.

**FINDINGS**

Planning for the digital surrogate bagging project began with biweekly meetings in April 2021 and continued through June 2021, with implementation running from July 2021 to February 2022. Staff packaged 68 digital collections ranging from 770 MB to 3.74 TB with accompanying metadata for a total of over 860,000 files at nearly 11 TB. Assessment procedures resulted in an investigation of 10 per cent of the surrogate collection, with repackaging required for only two collections (error rate: 3 per cent) and a data re-harvest of only four collections (error rate: 6 per cent).

Throughout the project, librarians in each area of specialisation shared their knowledge across organisational units. Table 1 shows the high-level knowledge (general context and tools) common to all staff, contrasted with the deeply specialised knowledge possessed by each unit. Sharing this specialised information among units, as well as within the knowledge areas identified as context, process and schema silos, proved to be an efficient model for collaboration on a complex project. The frequent, multidirectional exchanges of information also belied the notion of organisational silos as isolated groupings, separated by rigid organisational boundaries, that tend to hoard information and hinder cooperation. Rather, the organisational silos functioned as units of specialised knowledge, interconnected as needed by multiple communication channels.

Despite the efficient use of professional staff time, more time could have been spent on establishing common communication practices and gathering contextual information prior to implementation. During the early stages, it was not always clear who to ask contextual questions, which if asked earlier could have provided a more efficient outcome. Vantrappen and Wirz’s recommendation to use bridges built on shared values and networking could create valuable lines of communication and cultivate reliable knowledge-sharing behaviours.54

Additionally, a broad understanding of the greater context by all team members could help them to apply their respective expertise. For example, if a staff member was asked to add a specific element to a metadata set and they did not understand what its purpose was in the context of the project, then they would be less likely to suggest a better field that could be used for that purpose. In essence, their output would better represent that of an assembly-line worker than a co-innovator. An example of this occurred when one of the team members — who was not involved in the initial meetings — mistakenly assumed that some of the digital objects were new resources rather than surrogates that had been created several years earlier and were just now being accounted for in the MARC records. The Bertoldi et al. team’s respect for the ‘individualised needs’ of their human infrastructure will be an excellent model for cultivating this greater context as staff move forward with a system-wide approach to digital preservation.55
Finally, professional staff besides the digital preservation librarian could have been involved throughout the implementation phase to support review and assessment procedures. This could have minimised minor packaging errors. More importantly, collaborative review procedures would have given the intern more experience working in a culture with good models for collaboration and a supportive environment to make and learn from mistakes. Radical collaboration, as outlined by McGovern, could be a means of fostering a more inclusive work environment for all library workers.56

Overall, the team found that the subcategories of the PREMIS data model provided a helpful guide for identifying needs and implementing solutions, and organisational silos provided the deep expertise required to plan the project and review the products. The framework drew efficiently on the expertise of multiple areas of specialisation and supported the organisation of collaborative effort. The knowledge silos exposed through the framework allowed staff to build lines of communication for effective advocacy, process development and handoff procedures. The organisational silos proved integral to the successful planning of this project, and corroborated Vantrappen and Wirz’s observation that silos aggregate expertise, assign accountability and provide identity.57

CONCLUSION
Rather than barriers to break, silos can be reconsidered as valuable infrastructure in a dynamic library. Approaching silos in this way makes sense from a practical perspective, as they are likely to persist. It also acknowledges the expertise of diverse information professionals, as well as the necessity of specialisation for the organisation, maintenance and preservation of information artefacts. With the insights gained from the planning, execution and detailed reflection of this project, the authors are now better prepared to collaborate across silos and to maximise their potential to manage complexity.
Appendix: Questions asked during investigation

- **Bitstream entity questions**
  - Does the library have secure, reliable, backed up, fixity-monitored digital storage?
  - Who should have access to this storage?
  - Are there procedures in place for generating and verifying integrity?
  - What are the procedures for transferring into this environment?
  - What are the procedures for deletion?

- **Digital file entity questions**
  - What files are being managed?
  - What are the significant properties of these files?
  - What method of integrity capture is in place?
  - How will characterisation and validation be conducted?
  - How are the digital files related to their metadata records?
  - What packaging structure will be used?
  - What is an appropriate amount of contextual information to include?
  - What schema will be used to capture contextual information?
  - What content standards will be used?

- **Representation entity questions**
  - How are the files related to each other?
  - Where is data about their relationship captured?
  - How is this data structured?
  - Is the data structure important for preservation management?
  - Are there layers of relationships or complexity about these digital objects?
  - How much knowledge of past practice is necessary for understanding or using these files?

- **Intellectual entity questions**
  - How are they described at the collection vs item level?
  - How are these files understood as intellectual entities?
  - How are they discovered as intellectual entities?
  - Where are they inventoried?
  - What content standards are applicable?

- **What support files are available?** Will support files be carried forward?
- **Why are there extra TIFFs in the package?** Will these files be carried forward?
References


9. Kowalski, ref. 6 above.


15. Ibid.

16. Vantrappen and Wirtz, ref. 12 above.


18. Tett, ref. 1 above.

19. Vantrappen and Wirtz, ref. 12 above.

20. Henman, ref. 14 above.


26. Pryse, ref. 24 above.

27. Bertoldi and Wirtz, ref. 25 above.


31. MacDonald and vanDunkerken, ref. 11 above.

32. CCSDS, ref. 3 above.


38. McGovern, ref. 29 above.

39. Ibid.

40. Pryse, ref. 24 above.
41. McGovern, ref. 29 above.
42. Kowalski, ref. 6 above.
48. CCSDS, ref. 3 above.
49. PREMIS, ref. 4 above.
52. PREMIS, ref. 4 above.
54. Vantrappen and Wirtz, ref. 12 above.
55. Bertoldi, Griesinger and Narlock, ref. 25 above.
56. McGovern, ref. 29 above.
57. Vantrappen and Wirtz, ref. 12 above.