Building for Tomorrow

White Paper

Version 1.0

16 June 2021

Ann Whiteside (PI), Stephen Abrams, Sara Rogers, Kathlyn Kao and Hanan Kataw

This project was made possible in part by the Institute of Museum and Library Services LG-73-17-0004-17IMLS

The views, findings, conclusions or recommendations expressed in this publication do not necessarily represent those of the Institute of Museum and Library Services.

Contents

Acknowledgments 3

Executive Summary 4

Introduction 5

Problem Statement 5

Building for Tomorrow Grant Work 6

Recommendations 11

Appendices 12

Appendix 1: 2018 National Forum Summary 13

Appendix 2: Summary of Software Vendor Interviews (2019) 15

Appendix 3: Summary of Design Firm Interviews (2019) 17

Appendix 4: Summary Report of November 2020 Invitational Meeting 20

Appendix 5: Recommendations for Long-Term Preservation of Digital Design Data 25

Appendix 6: Transfer Guidelines and Resources 37

Appendix 7: Standards Resources 59

Appendix 8: Principles Resources 69

Acknowledgments to individuals who have contributed to Building for Tomorrow since 2019

John Bacus, Senior Director of Product Management, Design & Engineering Solutions Group at Trimble Inc.
Christopher Bavitz, Managing Director, Cyberlaw Clinic at the Berkman Klein Center for Internet and Society at Harvard University
Sara Bush, Archivist, Sasaki
Brandon Butler, University of Virginia, Director of Information Policy, University of Virginia
Luis Cetrangolo, Director of Digital Practice, Sasaki
Kyle Coburn, Senior Associate, Höweler + Yoon
Kyle Courtney, Copyright Advisor, Harvard University
Martien de Vletter, Associate Director Collection, Canadian Centre for Architecture
Nancy Hadley, Assoc. AIA, Director, Archives & Records, The American Institute of Architects
Eric Howeler, Höweler + Yoon and Associate Professor of Architecture, Harvard GSD
Leslie Johnston, Director of Digital Preservation, U.S. National Archives and Records Administration
Maggie LaPoint, Student (Spring 2021), Harvard Law School Clinic
Aliza Leventhal, Assistant Section Head of Technical Services for the Prints & Photographs Division of the Library of Congress
Moira O'Connell-Morganstein, Facilities Archivist Facility Information Systems, MIT
Jessica Quagliaroli, Architecture Records Archivist, Manuscripts & Archives, Yale University Library
Dan Rossiter, British Standards Institute
Dr. Katherine Skinner, Executive Director, Educopia, Inc.
Marek Suchocki, Sales Development Executive, Autodesk
Sylvia Welsh, Lead Archivist Property Information Resource Center, Harvard University

Executive Summary

The Frances Loeb Library at the Harvard University Graduate School of Design received an Institute of Museum and Library Services (IMLS) National Forum Grant (LG-73-17-0004-17I) under the National Digital Platform funding priority for Building for Tomorrow in 2017. The grant was to support two meetings of engaged stakeholders – architects, architectural historians, archivists, librarians, technologists, digital preservationists, and others to frame a collaborative infrastructure to support the long-term preservation of digital design data. With a grant extension from IMLS, we were able to continue work in 2019, engaging the CAD software vendor community through a series of interviews. The interviews covered questions related to the software development process, interoperability of the software, exportability of file types, ability to use free readers of the software, how vendors archive their software, and their views on the future of their software products. We also began a pilot project with the Sasaki design firm and the Höweler + Yoon architecture firm, both Boston-based firms, working with in-process design projects. During 2020 and 2021, with another extension from IMLS due to the COVID pandemic, the focus has been on developing preservation recommendations that address the preservation of the original files and the ability to experience those files appropriately to the software from which they are created, and investigation into the intellectual property issues in collections of design records, and the impact on eventual collecting institutions. This paper provides documentation of the grant, the outcomes, recommendations for preservation, addressing intellectual property issues, and recommended next steps for continuing work in this area of digital preservation.

Introduction

As the architectural and engineering fields rapidly embraced the application of information technology in all aspects of their activity beginning 50 or 60 years ago, museums and archives are now faced with the challenge of ensuring the preservation and future accessibility of the rapidly growing digital architectural artifacts, grappling with the need for technological tools, or how to adapt existing tools, technical and archival expertise, and repositories that can preserve and disseminate the archived digital data. The Building for Tomorrow grant has aimed to frame a collaborative infrastructure to support long-term preservation of digital design data. “Infrastructure” includes the technologies and tools required for digital archiving, preservation repositories, addressing intellectual property issues, and engaging the stakeholder communities – software vendors, designers, architects, architectural historians, archivists, librarians, technologists, and digital preservationists – in working towards shared understanding and agreement on how to best preserve and ensure accessibility to digital design records into the future.

Problem Statement

The earliest adopters of the new computer-aided design (CAD) software were industries like aerospace and automotive, and since then the fields of architecture and design have been enthusiastic adopters. While experimental use of digital computers in architecture can be traced back to the mid-twentieth century, they first entered large architectural offices in the U.S. in the 1960s through the “back door” of the engineering and management departments. In the following decades, the development of minicomputers and later personal computers was accompanied by an increasing number of commercially available CAD software packages. These early commercial CAD programs were predominantly 2D-based until the late 1980s when 3D solid modeling and parametric techniques became more available. All of which have resulted in a widespread use of computers for architectural drafting, modeling and, by the late 20th century, as creative design tools.

The use of 2D and 3D CAD and Building Information Modeling (BIM) software is now routine in architecture and design firms; the contractual deliverable is now usually an electronically signed model that can be manipulated to achieve equal or more granular information than print plans. The challenge is that the number of data files created for each design project can be enormous, the resulting data files become obsolete very quickly due to continuous upgrading of software. For designers, and for architectural archives and museums, the ability to preserve the data files and make them continually accessible and usable as originally created data is therefore a huge challenge.

Building for Tomorrow Grant Work

Building for Tomorrow, funded by the Institute of Museum and Library Services, IMLS Grant LG-73-17-0004-17IMLS, began in 2017. The Frances Loeb Library at the Harvard University Graduate School of Design received an IMLS National Forum Grant under the National Digital Platform funding priority to support two meetings of engaged stakeholders – architects, architectural historians, archivists, librarians, technologists, digital preservationists, and others to frame a collaborative infrastructure to support long-term preservation of digital design data. Building for Tomorrow is focused exclusively on born-digital files generated in the design process. In the case of digitized materials, well established preservation policies are already available and are being used in practice across libraries and museums. Work on standards for 3D scanning and scanned materials is advancing rapidly through the work of Community Standards for 3D data Preservation, which will be coming out with a monograph in 2021 through ACRL.

The goals of the forum, held in April 2018, were: (i) to set priorities for long-term preservation of digital design documents; (ii) to think collaboratively about issues in preserving architectural design data; (iii) to find alignments of needs and challenges across communities; (iv) to identify a path of development for sustainable shared infrastructure for preservation of digital design documents and data, usable by a variety of types and sizes of architectural museums and archives, in the form of an action plan for the next three to five years. An infrastructure in this context includes human resources required to do this work, along with the technological tools, methodologies, and services needed to support institutions of varying types and sizes. The Forum resulted in an extensive plan of action for work to move forward.

With a grant extension from IMLS, we were able to continue work in 2019, engaging the CAD software vendor community through a series of interviews. The interviews covered questions related to the development process of software, interoperability of the software, exportability of file types, ability to use free readers of the software, how vendors archive their software, and their views on the future of their software products.
This provided useful information about shifts in the software industry that impact the work in the design professions and the outlines for a collaborative pilot project with two design firms. In Fall 2019, we began a pilot project with the Sasaki design firm and the Höweler + Yoon architecture firm, both Boston-based design firms. The pilot work began with meetings with the designers for two projects at Sasaki, and the designers for one project at Höweler + Yoon. We talked about software use through the design process, different file types created and used, how files are organized within projects and within the firm overall, what files they keep as part of their archives, when handoffs happen, what marks specific handoffs, and intellectual property issues. The intent was to work with the firms to ultimately archive the specific projects through the design lifecycle, and unfortunately, the COVID pandemic stopped us from proceeding.

During 2020 and 2021, with another extension from IMLS due to the COVID pandemic, the focus has been on developing preservation recommendations that

address the preservation of the original files and the ability to experience those files appropriately to the software from which they are created, and investigation into the intellectual property issues in collections of design records, and the impact on eventual collecting institutions. We engaged the broad spectrum of stakeholder communities in this work with us through virtual meetings in November 2020 and through the spring of 2021 with subsets of the stakeholder groups. We engaged stakeholders in discussions about the intellectual property issues in the winter and spring of 2021.

The preservation recommendations meetings in November 2020 resulted in the formation of four sub-groups that focused on work that would support the emerging preservation recommendations.

One sub-group focused on identifying the ISO and other informal standards that support the work of designers through the design lifecycle and have impact on the archiving of the digital data later. The result was a near comprehensive list of standards that can be referenced from the viewpoints of the custodial institution, the designer or creator of the original data files, and the software vendors. The recommendations for next steps in this paper identify further work that can be done in this area. The standards work is documented in this report and on the Building for Tomorrow website.

A second sub-group worked to establish a set of transfer guidelines and resources to facilitate preparation of content for transfer from a designer to an archive. The resources are included in this report and are available on the project website.

A third sub-group worked on developing a resource list of recommended principles to guide implementation and document a spectrum of practices and software tools and platforms that may satisfy those principles including for small or low-resourced archives, federal or other very specific/regimented archives and broader, more highly-resourced archives. The Principles are found in this report with the Standards documentation and on the project website.

A fourth sub-group was identified to focus on access with a goal to develop a shared set of requirements for access, including requirements based on real life scenarios of what researchers have requested or used in the past. Due to time constraints, work on the area of access to preserved content was not pursued during 2021 and is identified as a topic that requires further work.

In October, 2020, the project team met with the American Institute of Architects Archivist Nancy Hadley and AIA legal representative Michael Bomba to discuss questions about how we can work with the AIA to develop more direct ways to inform designers about the importance of thinking about intellectual property issues beyond the moment of a contract with a client, and with an eye to thinking about the future disposition of their work. The AIA offers a suite of contracts including addenda and

we were offered the option of developing an addenda to be included in the AIA suite of contracts. In January, 2021, the project team met with Christopher Bavitz at the Harvard Law School Cyberlaw Clinic, Brandon Butler, Director of Information Policy, University of Virginia, and Kyle Courtney, Copyright Advisor, Harvard University to discuss the legal issues surrounding intellectual property of digital design files.

Following that meeting, in collaboration with the Harvard Law School Cyberlaw Clinic, the Building for Tomorrow team worked during the spring of 2021 to begin development of a collection of template agreements that would facilitate contribution of architectural materials to libraries and archives in a format that enables dissemination to and access by the research community. These took two primary forms.

First, Building for Tomorrow and the Clinic reviewed examples of standard form donor agreements often used by libraries and archives and evaluated how well they suit the realities of a world in which digital access is desired and expected by both institutions and the research community. Unsurprisingly, traditional donor agreements contemplate contribution of physical materials and on-site access by researchers (along with attendant temporal and other limitations to such access). Donation of physical drawings and related materials implicates intellectual property law minimally if at all, while allowing for online access requires consideration of copyright protections and accomplishes the requisite grant of rights.

Specifically, crafting a donor agreement that reflects an interest in making materials available online requires consideration of:

(a) which parties have the copyrights or other related rights necessary to grant a library or archive the ability to make materials available online; and

(b) the scope of any grant of rights, including but not limited to whether rights should be assigned or licensed to the institution.

Regarding the first of those two points — any effort to modernize donor agreements will likely have to be accompanied by an educational effort directed to staff of libraries or archives in order to evaluate given rights situations and position them for success when engaging with donors, bringing in new collections, and establishing terms on which those collections will be made widely available.

Second, Building for Tomorrow and the Clinic investigated the possibility of developing standardized contractual language that might be added to or incorporated in agreements between architects and builders at the earliest stage of a project, setting out the parties' intention to contribute materials to an archive from the outset or at some future date. By declaring that intention early on, Building for Tomorrow believes that builders, architects, and archival institutions might avoid questions that would otherwise arise later about the ownership of rights or the scope of rights granted. Further, by encouraging parties to consider and address these issues from the outset of a project,

Building for Tomorrow believes that the culture and norms in the field could develop over time to promote a baseline expectation of availability and access.

This work has taken the form of proposed contractual language that might accompany standard agreements promulgated by the American Institute of Architects. The proposed language can be crafted to address various rights ownership and licensing scenarios that might characterize the full range of owner-architect arrangements. Conversations are underway with AIA and others in the field about refining this language and developing mechanisms to make it available and educating the broader community about its applicability.

Harvard is also participating in a pilot with Emulation as a Service (EaaSI) and the Software Preservation Network (SPN) in a one-year pilot of the EaaSI beta hosted service for participants in SPN member organizations. The goals of the pilot are to determine ways of standardizing and scaling software and environment, configuration of workflows across participating SPN member organizations, and to determine the level of SPN staff support and associated costs that would be required to make an ongoing partnership with EaaSI feasible. From the perspective of Building for Tomorrow, this allows the Harvard team to test EaaSI as one of the tools we need as part of our digital preservation toolkit, and provide access to design files for researchers. As of June, 2021, the pilot is in early stages. We expect to have information to report at the end of the 2021 calendar year.

Parallel Work

Parallel work in this space has been going on as well, which has been helpful to Building for Tomorrow, and has allowed the team to identify areas of collaboration going forward. The Society of American Archivists Design Records Section has been doing related research and surveys in the archival community since 2012. The 2014 International Confederation of Architecture Museums (ICAM) conference addressed these issues in several presentations. In November 2017, the Library of Congress, the Architect of the Capitol, and the National Gallery of Art hosted a summit, Designing the Future Landscape: Digital Architecture, Design and Engineering Assets, at the Library of Congress. Finally, the Canadian Centre for Architecture has led the effort for the archival profession, putting into practice the use of tools and techniques from the digital preservation community.

The Society of American Archivists Architectural Records Roundtable launched a CAD/BIM Task Force in 2013 as a catalyst for a community-wide initiative to address the numerous (legal, technical, and curatorial) issues of born-digital architectural records. The work of the task force includes a survey of firms and architectural archives to learn about holdings and current archival practices for born digital design data, a Born Digital Studies Bibliography, and a report of the task force. The CAD/BIM Task Force has changed its name to the Digital Design Records Task Force and has produced several important documents to support the work of archiving digital design records. These include surveys,

information from interviews with designers and archivists, an SAA Research Forum Paper, and an Archival Description project and deed of gift matrix.

In November 2017, the Library of Congress, the Architect of the Capitol, and the National Gallery of Art hosted a summit titled Designing the Future Landscape: Digital Architecture, Design and Engineering Assets, at the Library of Congress (Leventhal 2018). The summit brought together 180 stakeholders in the architecture, design, and engineering professions, from creators to curators, to explore the issues and obstacles of long-term preservation and access to the records of their projects and to begin working toward sustainable solutions. Critical issues that arose through the summit were: the need to identify the full array of digital design files created in1, the need to determine which design records and specific types of data or information in those records the various stakeholders need in the immediate and long-term future, and the need to develop better communication and information-sharing practices, which are critical to developing sustainable solutions to problems with the long-term preservation, access, and use of digital design files. Having so many representatives from the range of digital design communities engaged in discussion made clear that the issues of preservation have impact across all stakeholder groups, as well as the risk we face in losing cultural heritage information. Discussions highlighted the fact that advancements in digital preservation tools, such as BitCurator and Archivematica, and collection management tools such as ArchivesSpace, offer technological support for the preservation of digital design files. The communities at the summit also recognized the importance of working with colleagues across domains to support the preservation efforts.

Recommendations

- Disseminate the work of Building for Tomorrow through:
  - Publication of this White Paper
  - Building for Tomorrow website
  - Through article and conferences
- Intellectual Property work:
  - Share the proposed draft contract with the American Institute of Architects for feedback.
  - Share the proposed draft contract with other stakeholders including designers and archivists.
  - Share draft donor agreement template that can be used between donors of archival collections to archival institutions with stakeholders including designers and archivists and librarians.
  - If the AIA agrees to incorporate contract language or an addendum to their suite of contracts, work with the AIA on educational programming around contracts.
- Standards:
  - Provide additional guidance on how to apply the standards recommended by other standards bodies, such as standards bodies located in the UK and Europe
  - Investigate BIM standards work in Europe and the US and provide guidance on how to apply emerging standards and identify impacts on archiving institutions.
- Continue work with the SAA Digital Design Records Taskforce
  - Share information with digital preservation software developers related to the creation and use of digital design data in order to ensure future product development will include infrastructure to support the long-term preservation and access of digital design files.
  - Collaborate on efforts to continue to bring design practitioners, managers responsible for a firm's technology strategy and documentation, and software developers to the table with archivists and digital preservationists.
- Providing access to digital design files:
  - Develop a shared set of requirements for access, including requirements based on real life scenarios of what researchers have requested or used in the past.
- Educating Design Students in design file data management:
  - In collaboration with faculty in design schools, develop a data management training curriculum focused specifically on design records.

Appendices

Appendix 1: 2018 National Forum Summary

Appendix 2: Summary of Software Vendor Interviews (2019)

Appendix 3: Summary of Design Firm Interviews (2019)

Appendix 4: Summary Report of November 2020 Invitational Meeting

Appendix 5: Recommendations for Long-Term Preservation of Digital Design Data

Appendix 6: Transfer Guidelines and Resources

Appendix 7: Standards Resources

Appendix 8: Principles Resources

Appendix 1: 2018 National Forum Summary

The National Forum in 2018 brought architects, architectural historians, archivists, librarians, technologists, digital preservationists, and all those engaged in preservation of digital design data together for a day and a half-long workshop. Thirty-five participants, invited to represent these areas of expertise were asked to work together in small and large groups to meet the desired outcomes of the Forum. The goals of the forum were: (i) to set priorities for long-term preservation of digital design documents; (ii) to think collaboratively about issues in preserving architectural design data; (iii) to find alignments of needs and challenges across communities; (iv) to identify a path of development for sustainable shared infrastructure for preservation of digital design documents and data, usable by a variety of types and sizes of architectural museums and archives, in the form of an action plan for the next 3-5 years. An infrastructure in this context includes human resources required to do this work, along with the technological tools, methodologies, and services needed to support institutions of varying types and sizes.

The forum was held immediately prior to the Society of Architectural Historians annual conference on April 17-18, 2018, in St. Paul, Minneapolis. It was led by Building for Tomorrow PI, Ann Whiteside, and facilitated by Christina Drummond (consultant) with help from Aliza Leventhal (Library of Congress), Jessica Meyerson (Educopia), Kit Arrington (Library of Congress), Kari Smith (MIT), and Kate Neptune (Harvard Business School) to keep discussion on topic, to encourage all voices to be heard, and to prompt critical reflection within small groups.

The goals of day one were to identify barriers, issues, and needs of the different stakeholder groups represented; to identify priorities in developing a shared "infrastructure"; and to map out connections between and among the communities represented at the Forum and other communities with which we need to develop partnerships. The end product of the day was a working "map" that included collecting stages of design files, the work flows for each area of expertise, and capturing the multi-dimensional perspectives from each area of expertise (for example the architect is concerned with different parts of the Architecture, Design, and Engineering (ADE) record lifecycle than the archivist) of preserving born digital design files and the various connections to other groups and organizations that are working on similar issues with which we need to connect. This "map" visually described the work that needs to be accomplished to develop infrastructure to support the preservation of digital design files from creation through ingest by collecting institutions.

Day two focused on moving from the mind map created on day one towards developing a set of strategic directions and actions for a community driven infrastructure for preserving digital ADE records. During this day a clarification to the moniker for ADE records was proposed: DADE (Digital Architectural Design Engineering records). This was to clarify the project's orientation towards new (post 2010) digital ADE records: works that are born-digital within the lifespan of current software versions, rather than legacy works

that will benefit from solutions for new digital records but require additional technologies such as emulation.

A follow up meeting was held in late May of 2018, with a subset of the Forum attendees to work on refining and finalizing the strategic directions. A high-level view of the strategic directions is provided here. The full document can be found on the Building for Tomorrow site.

The Strategic Directions:

1. Understanding and Representing the Current State of the Field

The goal of this Priority is to understand the landscape of digital preservation efforts surrounding the DADE coalition, how the efforts are aligned or in dialogue with each other and DADE, and to share this perspective with members of the DADE community in order to identify opportunities to better support the community and to prioritize the work of the committee.

2. Start and Improve Coordination across Institutions on DADE and Connect to Stakeholders

The goal of this field is to develop community engagement means and methods.

3. DADE Coalition Governance & Sustainability

The goal of this item is to make the DADE coalition sustainable in the long-term through clear organization, strong commitments of community members and continued pursuit of funding.

4. Improve the DADE community's ability and capacity to preserve DADE records

This item catches the known gaps and services needed within the DADE community.

Appendix 2: Summary of Software Vendor Interviews (2019)

In summer 2019, Sara Rogers and Ann Whiteside interviewed representatives from four CAD software firms (Autodesk, VectorWorks, Trimble, Bluebeam). The questions we brought to each representative were the following:

1. What are the major milestone moments in the evolution of a software or the software suite?

    a. And in also taking technological breakthroughs and changes in computing so else the company's approach and intention in what the software could do.

2. What was the build up like between each milestone?

    a. How forwards and backwards compatible are those files or the features influenced by user requests or solely felt in product roadmap?

3. What leads to changes in the interfaces?

    a. How do you think about how the user experience is impacted?

4. Is software intentionally inter-operable or engaged with other vendor softwares?

    a. Are you intentionally working with other vendors and if so who they are?

    b. What file types are made of, supported, importable or exportable by your software?

    c. How are updates pushed out?

    d. Does the company themselves keep copies of all the version of their softwares?

    e. And if yes, is the software functional? Is the thing maintained?

5. Would you be willing to share preservation or access copies to collecting institutions that have files made with these versions of their softwares?

6. What is the best free reader of the software?

    a. What are the missing features from the free reader versions?

    b. Do you feel that they are valid?

7. What do you see as the most critical view or the record of the files that your software produces?

We learned that in general, the focus on developing software is on continuous improvement and iterative changes, based on the needs of the software users. The compatibility of software is generally forward, though most vendors have limited backward compatibility. The focus of forward compatibility is driven by software sales to users who are looking for the latest versions of software. Each of these has implications for archiving institutions because the software used in a design is generally not accessible after a period of time and is a known issue for preservation. Focus is on continuous improvement or iterative changes to software to accommodate changing needs of software users.

Two vendors shared their approach to a product roadmap; other were less specific in describing their approaches.

Overall, we hear repeatedly that the future is about the use of Google documents and the Cloud. They allow much more sharing of information across project participants in dispersed locations.

There is also a focus on the integration or compatibility of multiple software tools so that the user doesn't notice that they are using different tools together. The user no longer has to switch in and out of different software programs when the software is integrated.

The vendors themselves indicated that PDF is the most critical view of the files produced with the software, and the one that they would rely on as a record of the software.

Appendix 3: Summary of Design Firm Interviews (2019)

The Harvard Team (Stephen Abrams, Sara Rogers, Ann Whiteside) met with designers and a BIM manager in two firms – Sasaki (a large, international firm) and Höweler + Yoon (a small Boston-based firm) in fall 2019 and winter 2020. Sasaki offered meetings with two different design teams working on two separate projects, providing us with projects of different scope and size, each of which was in-process. Our meetings with Höweler + Yoon were held with one of the principals and a partner and focused on one in-process project.

The questions we brought to each design team were the following:

a. Questions:

i. Where are the handoffs?

1. Design development phase
2. Schematic design
3. Construction documents
4. As-builts
5. Others?

ii. What is the journey of the plan? Going back and forth between?

iii. How are the handoffs happening?

6. Email
7. WeTransfer

iv. What do designers view as the most important document?

v. What is the evolution of the document?

vi. How do you close out projects?

vii. How are software products implemented and used?

8. If you didn't have X software what would you do?

The themes that emerged from the interviews are:

Use of software products:

- Since about 2008, all software that is BIM related is not perpetual anymore. Software is paid by subscription, so there is no physical software to go back to. There is a three-year window with updates.

- AutoCAD is used primarily as a documentation tool in architecture, though still heavily used as a design tool in landscape architecture. And is becoming obsolete because it does not always translate in the same way when shared with other software

- The industry is moving to Revit as the software of choice because it is easier to manage works more easily with other software

- There are numerous software products used throughout the design life cycle, including Rhino, Revit, SketchUp, and others, depending on the firm

- A wide range of software is used for presentations: InDesign, PowerPoint, bitmap, jpeg, tiff; (presentation documents) InDesign, PDF, ppt; Hand sketches (jpeg); computer generated (renderings) photoshop, jpeg, tiff; (line drawings) illustrator, AutoCAD, Rhino

- Other tools such as Excel are incorporated for different purposes in a project

- "Record" drawings that eventually become "as-builts" that don't usually come back to a firm

- The most critical to save are early design ideas and the final BIM, although having as much as possible throughout the whole process would be highly desirable

- Early sketches are often still hand-drawn, which are not always scanned

- Each project will have unique workflows based on the needs of the project and the tools available

- Digital Artifacts *design documentation* – (3D design model) rhino, (3D rendering model) 3DS Max, (BIM) Revit, (construction) PDF

File management:

- Models are saved at key milestones, typically at design preservation

- File structures are often project-driven

Preservation of files on the firm side:

- Most companies are very concerned about not losing work doing in the past 2-3 days. If something breaks, there is a need to have the resources to recover what was lost. Weekly backups, already served in our servers or colocation. If there is a data loss disaster, there is the option of hiring consultants to recover the data.

- On the side of IT and legal implications related to preservation, data is usually kept for three to six months, and extending back to a year. One firm noted that they now have constant back up kept locally and in a co-location. Generally, everything is backed up because it takes more people hours to select what to backup than to save everything.

- Legally, the firm is required to retain record documents for 15 years and there is not time to filter through all the files. Across the board, because of state legal requirements (in Massachusetts), international work has to comply with higher legal standards.

- Software vendors do not want to commit themselves to preservation. There could be greater weight to the issue if the architectural community could go back to the vendors and talk to them. Sasaki is willing to help collaborate on this, or on the prospect for more widespread understanding in architecture.

Building Information Modeling (BIM):

The use of BIM is for automating access to files from stakeholders involved in the design process to construction, and through to the ultimate building managers once the building is complete. BIM is a paradigm shift in the design process because it is not a tool per se, but a database that facilitates process across disciplines. When each stakeholder of project needs something or needs to work with something, idea is that you have access to certain aspects of the database. In the database, you have schedules of lists of different things, the stakeholder will extract the list of components into their work, and design and construction changes are shared within the BIM automatically.

Education of design school students in data management:

As part of the interview process, we also discussed design school pedagogy and future education of design students. We consistently heard from all the designers we interviewed that graduates from design schools enter firms with very little understanding of data management in the context of design practice, and that BIM is not taught in all schools. The lack of education in these areas means that firms have to develop very rigorous training in these areas as they onboard new staff. This is noted as an area for future work on the part of design schools and librarians and archivists in the schools.

Appendix 4: Summary Report of November 2020 Invitational Meeting

The objectives of the meeting were to:

- Identify key impediments to the effective and sustainable long-term preservation of digital architectural and design assets and offer concrete responsive recommendations for best practices

- Strengthen understandings between stakeholders regarding their different perspectives and their united goals

- Identify and refine concrete activities these stakeholders will engage in to forward general preservation aims and these specific recommendations

The outcomes of the meeting were:

- To achieve agreement on principles for preservation of digital architectural and design assets, including a clear focus on the preservation of meaningful *experiences* of architectural design information, which implicitly encompasses concern for both authentic digital assets and the intermediating software necessary for use and reuse of those assets.

- To develop an action plan for encouraging and supporting adoption across the stakeholder communities, which includes designers, archivists/librarians/curators, digital preservationists, emulation experts, faculty, design firms, historians, and organizational representatives.

Attendees:

- Seth Anderson, Software Preservation Program Manager Preservation: Digital Preservation, Yale University
- John Bacus, Senior Director of Product Management, Design & Engineering Solutions Group at Trimble Inc.
- Daniel Cardoso-Llach, Associate Professor, Carnegie Mellon University
- Euan Cochrane, Digital Preservation Manager Preservation: Digital Preservation, Yale University
- Martien de Vletter, Associate Director Collection, Canadian Centre for Architecture
- Nancy Hadley, Assoc. AIA, Director, Archives & Records, The American Institute of Architects
- Eric Howeler, Höweler + Yoon and Associate Professor of Architecture, Harvard GSD

- Leslie Johnston, Director of Digital Preservation, U.S. National Archives and Records Administration
- Aliza Leventhal, Head of Technical Services Section, Prints & Photographs Division at Library of Congress
- Clifford Lynch, Executive Director, Coalition for Networked Information
- Moira O'Connell-Morganstein, Facilities Archivist Facility Information Systems, MIT
- Emily Pugh, Digital Humanities Specialist for The Getty Research Institute
- Marek Suchocki, Autodesk
- Tessa Walsh, Digital preservationist and software developer, BitCurator

The meeting was facilitated by Dr. Katherine Skinner, *Executive Director, Educopia Institute*

Building for Tomorrow project team

- Ann Whiteside (PI), Assistant Dean for Information Services, Harvard University Graduate School of Design (GSD)
- Sara Rogers, Digital Archivist, GSD
- Kathlyn Kao, Graduate Student, GSD
- Hanan Kataw, Doctoral Student, GSD
- Stephen Abrams, Head of Digital Preservation, Harvard Library

The agenda:

The meeting took place over the course of two half-days.

Day one:

- Introductions
- Overview of preservation challenges stakeholders face
- Workflow design and troubleshooting
- Recommendations for action

Day two:

- Realizing the collective benefits of our work
- Recommendations and Actions development
- Commitments

Themes that arose in day one included:

- A lack of shared nomenclature between stakeholder communities (I.e., “design, bid build” as opposed to “design and construction”)

- A tension between the technologist presumption that the future will take care of itself and the archivist presumption that complexity will continue to abound.

- There are existing standards and guidance documents that no one knows about or implements because they are used in different communities arose multiple times; solving this requires identifying those standards and then serving as emissaries

- One of the most important conversations was about how decisions are made and who makes them about what projects and/or components of projects are archived. What roles does privilege play in the decision-making today, and how can we push against that, or at least make it more visible?

- The need to work collaboratively across stakeholder communities is clear, and at the same time, individuals have bandwidth constraints. How to move this work forward in the future will require deep collaborations across communities, and very clearly delineated and discrete tasks for each step.

The recommendations that arose out of the first day's work were:

1. To develop good practice documentation

    a. Action: Establish a set of recommended principles to guide implementation and document a spectrum of practices that may satisfy those principles including for small or low-resourced archives, federal or other very specific/regimented archives, and broader and higher-resourced archives

        i. Principles list ("Need to know what you have" "Need to avoid single points of failure" etc) that builds from existing sources (eg CCA, NARA, NDSA) but frames the principles with a specific design focus and exemplifies the spectrum of practices

        ii. Collections considerations documentation or framework that highlights what collection gaps are known and need to be addressed (eg, underrepresented groups)

    b. Action: Develop a shared set of requirements for transfer to an archive (not building a tool) but just agreeing on what a complete submission information package looks like.

iii. This would give at least a starting point for changing recordkeeping practices earlier in the design/build/bid stage and also help archivists/practitioners state their needs for archival transfers.

c. Action: Develop good practice documentation for unstructured data (e.g., emails) building on existing documentation built for other subfields/domains

2. Develop a shared set of requirements for access - this should include requirements based on real life scenarios of what researchers have requested or used in the past. Again, whatever happens during processing will vary, but at least knowing what is needed as a baseline dissemination information package will be a starting point for figuring out what all happens during processing on a collection-by-collection basis

3. Engage with current/emerging standards and bring all stakeholder POV to bear on their implementation

a. Action: Collect existing standards and create an annotated list of standards currently in use, and note what voices are involved in each and what voices are missing from each (e.g., ISO, SAA, etc)

b. Action: Based on the missing voices, approach standards bodies with a request for engagement of additional perspectives

4. Develop initial vendor agreements (or a framework for these) to model how vendors and archives may collaborate towards shared aims

a. Action: EaaSI and Sketchup begin drafting an initial agreement; note challenges or issues that need to be solved

b. Action: Build a framework for collaboration that includes a list of who needs to be involved (positions), what questions need to be researched, general themes to address

iv. Test the framework, eg with a pilot by AutoDesk and CCA

5. Build partnerships with Design Schools to develop and pilot a simple initial set of training modules for students on how to structure design content for eventual hand-offs and longevity

a. Action: Develop a one-page advocacy statement and circulate to Design Schools to engage a broader set of instructors for future work

b. Action: Establish initial set of competencies for design students

v. Identify existing training opportunities, including in adjacent fields

vi. Identify training modules and frameworks that might serve as starting points

vii. Identify key stumbling blocks to long term preservation that should be included in the training (for example, issues around links to external libraries or between drawing sets)

Day two began with a discussion about what this collective of stakeholders is, exactly, and what kinds of work it can do together, what it cannot do, enabled everyone to think about their own role within the collective, and allowed us to work collaboratively to flesh out and refine the action items, and to define how the work would be done over the course of spring 2021.

Of the five actions agreed upon, we focused on four: good practice documentation (1) focusing on principles (1.A.i) and developing a shared set of requirements for transfer to an archive (1.B.); shared requirements for access (2); and engaging with current and emerging standards (3.A.).

Draft refinements to the charge were completed, and workflows brainstorms were completed. The workflows were finalized by Sara Rogers (Harvard GSD) after the meeting.

We reached agreement from most participants to continue work on these action items, with three to five participants for each of the four workgroups, and work began in January 2021, and was completed in May, 2021. The results of the work are found in this White Paper.

Appendix 5: Recommendations for Long-Term Preservation of Digital Design Data

Executive Summary

Effective preservation of digital design data is dependent on curatorial, technical, and legal capabilities and commitments. These recommendations provide guidance on all aspects consistent with general long-standing archival practices as well as the specific practitioner and academic needs, goals, and aspirations of the design and cultural memory communities. Conformance to these recommendations provides greater confidence in the successful and sustainable long-term preservation of digital design data.

1 Introduction

The architecture, engineering, and facilities management fields have embraced the application of information technology in all aspects of their activity. Ensuring the ongoing viability and accessibility of the digital design artifacts underpinning that activity is crucial to historical understanding of the built environment, effective education of future generations of scholars and practitioners, productive reuse and repurposing of the architectural record, and responsible operation and maintenance of realized architectural projects.

1.1 Purpose and Scope

This report makes specific recommendations for best practices to the architectural community of scholars and practitioners regarding effective, efficient, and sustainable preservation of valuable and often unique digital data.

For purposes of these recommendations, digital preservation is defined as the programmatic enterprise enabling meaningful human communication of digitally-encoded information across time and ever-accumulating technical and cultural distance. Digital preservation activity entails a complex of people, policies, procedures, and systems ensuring the archival integrity, authenticity, accessibility, and usability of proactively-managed digital resources.

Although the title of this document references “design data”, this language should be interpreted expansively to encompass all digital artifacts significant to exploratory, detailed, and final architectural, landscape, and engineering design, construction, and facilities management. The phrases “architecture, engineering, and construction” (AEC) and “architecture, design, and engineering” (ADE) are more descriptive of the full range of concerns targeted by these recommendations, however, that terminology is less familiar to the non-specialist audience and is not used here.

Development of these recommendations occurred in the context of Harvard University’s *Building for Tomorrow* (BfT) project

(https://projects.iq.harvard.edu/buildingtomorrow/home), generously funded by the US Institute of Museum and Library Services (IMLS), #LG-73-17-0004-17. The final recommendations incorporate critical perspectives of a wide range of consulted stakeholders as well as the participants of an invitational workshop organized as part of BfT project activity in November 2020.

1.2 Convention

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2019 (Bradner, 1997) when, and only when, they appear in all capitals, as shown here.

2 Prior Work

The importance of the need for effective and sustainable processes for protecting the long-term authenticity and usability of architectural design records has been recognized by the archival and practitioner communities for some time. These recommendations build upon a prior legacy of investigation of relevant issues, particularly with regard to these initiatives:

1. FACADE project (Future-proofing Architectural ComputerAided Design, MIT, 2007-2009)
2. LC Architecture Design and Engineering summit (Library of Congress, 2017)
3. Building for Tomorrow 2018 Forum

3 Objectives and Outcomes

The following recommendations define a set of baseline best practices for effective and sustainable long-term preservation of digital design data. The recommendations were developed through widespread consultation regarding the following core questions:

1. What should the archival and stakeholder design communities ***do*** (with respect to the long-term preservation of the design documentary record)?
2. What do they have the ***capability*** and ***capacity*** to do? (Capability refers to possession of the expertise to know *how* to do something; and capacity, access to the tangible resources -- staff, funding, and infrastructure -- necessary to *actually* do it.)
3. What do they have the ***legal*** and ***ethical right*** *to do?*

4 Recommendations

The following recommendations provide general guidance on best practices ensuring the long-term preservation of digital architectural design data.

R1 Overall practices MUST adhere to general archival principles and guidance as specified by ISO 15489 *Records Management* (ISO, 2016), digital preservation programmatic and system responsibilities as specified by ISO 14721, *Open Archival Information System Reference Model* (OAIS) (ISO, 2012a), and institutional and procedural trustworthiness as specified by ISO 16363, *Audit and Certification of Trustworthy Digital Repositories* (TDR) (ISO, 2012b).

R2 Specific policies, staffing roles, technical resources, and workflow procedures implementing the local application of these recommendations SHOULD be fully and openly documented by architectural archiving programs, including the scope and justification of any local variations, additions, or omissions.

R3 Local archiving analysis and decision-making MUST be cognizant of the needs, goals, and aspirations of various stakeholder communities designated as the target audiences of the archiving effort.

R4 Local archiving analysis and decision-making MUST be reviewed periodically to ensure continued relevance and accountability to existing, evolving, and newly-identified designated stakeholder communities.

R5 Archival programs SHOULD allow for curatorial designation of reasonable time-limited embargoes on public access to archived data to permit rights holders a legitimate but bounded time period in which to exercise sole intellectual and commercial exploitation of the data. These embargoes on public access to data SHOULD NOT apply to core descriptive metadata pertinent to high-level discovery.

4.1 Data

The following recommendations are specific to considerations of data formats, metadata, and packaging.

R6 For purposes of these recommendations, "data" MUST be interpreted expansively and inclusive of any visual or textual artifacts pertinent to the intellectual, aesthetic, and pragmatic conception, design, construction, and maintenance of architectural projects.

R7 Prior to transfer of data to archival custodial care, it SHOULD be managed locally in accordance with best professional IT data management practices, such as those recommended by the ITIL (AXELOS, 2019) and COBIT (ISACA, 2020) standards, including audited security enforcement and regularly-scheduled and verified backup, to ensure that authentic data are available at the appropriate point for their archival acquisition.

R8 Architectural practitioners SHOULD establish and conform to standardized practices for the structural organization and naming of data files in local IT infrastructure.

R8.1 Folder and file names SHOULD be self-descriptive, including delimited elements for architect name, project name, data function, version, and date.

R8.2 Naming standards MUST reflect the deliberate balancing of brevity and descriptiveness; established nicknames, initialisms, and acronyms MAY be used to minimize name length.

R8.3 Names SHOULD rely only on a file-system-safe subset of printable ASCII: 0-9, A-Z, a-z, and the period (or full stop), dash (or hyphen), and underscore:

0 1 2 3 4 5 6 7 8 9
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
a b c d e f g h i j k l m n o p q r s t u v w x y z
. - _

R8.4 File extensions MUST accurately reflect the underlying format of the data file.

R9 Data SHOULD be transferred to archival custodial care at various key milestones in the architectural information lifecycle, potentially including but not limited to:

1. Initial statement of design program/creative brief/client requirements/RFP/etc.
2. Initial design parti
3. Major conceptual phases of iterative design development
4. Final submitted project bid
5. Major conceptual phases of iterative design development
6. Construction bid
7. As-built

R10 Data SHOULD be captured, transferred to archival custodial care, and maintained in its original native form.

R11 Data SHOULD be additionally represented in the widest range of supplemental derivative forms inherently amenable to sustainable long-term stewardship (see Section 4.1.1).

R12 Supplemental derivative representations SHOULD be created and verified by the originating design agent(s) or, if necessary, at the point of contribution to or acceptance by an archiving program.

R13 Data SHOULD be accompanied by comprehensive metadata documenting the intellectual, aesthetic, engineering, administrative, technical, and relational context of the data (see Section § 4.1.2).

4.1.1 Format

The following recommendations are specific to considerations of data format.

R14 Selection of formats SHOULD maximize preservation sustainability in terms of factors identified by the US Library of Congress's *Format Sustainability* framework (Library of Congress, 2017) and minimize risk in terms of factors identified by US National Archives

and Records Administration's *Framework for Risk Assessment and Preservation Planning* (NARA, 2020).

R15 Consistent with these sustainability and risk factors, format selection SHOULD, in the absence of other compelling criteria or imperatives favor:

- Standardized, rather than non-standardized formats;
- Openly documented, rather than closed formats;
- Community developed/supported, rather than proprietary formats;
- Mature, rather than novel formats;
- Text, rather than binary formats;
- Legally unencumbered, rather than encumbered formats;
- Formats in common, rather than sparse use across the design community; and
- Formats supported by commodity, rather than specialized software tools.

R16 Selection of formats for creation of derivatives of native data representations SHOULD conform to these genre-specific recommendations:

- CAD:
  - Preferred: Autodesk Exchange File (ADSK), Design Web Format (DWF), PDF/E (ISO 24517), STEP (ISO 10303)
  - Acceptable: AutoCAD Drawing Database (DWG), Drawing Exchange Format (DXF)
- Still image:
  - Preferred: JPEG 2000 (ISO 15444) with lossless compression and embedded color profile
  - Acceptable: DNG, JPEG/JFIF, TIFF (uncompressed)
- Moving image:
  - Preferred: MPEG-4 (H.264), MXF
  - Acceptable: FFV1 in Matroksa
- 3D/AR/VR:
  - Preferred: PRV (ISO 14739-1), X3D (ISO/IEC 19775-1)
  - Acceptable: U3D (ECMA-363)
- Document/presentation:
  - Preferred: PDF/A (ISO 19005-4)
  - Acceptable: OpenDocument (ISO/IEC 26300), PDF (ISO 32000), Powerpoint (.pptx), Word (.xlsx)
- Text:
  - Preferred: Unicode in UTF-8 encoding and no byte order mark, ASCII (ISO 646)
  - Acceptable: ISO/IEC 10646 in UTF-8 encoding
- Data:
  - Preferred: CSV, TSV
  - Acceptable: Excel (.xlxs)

R16.1 All formats supporting the option SHOULD include embedded metadata documenting the derivative's legal status and terms of use.

R16.2 All formats supporting the option SHOULD include embedded metadata documenting the derivative's pertinent technical characteristics, e.g., XMP (ISO 16684-1), EXIF for still images (CIPA, 2019), etc.

R16.3 Still image files SHOULD include an embedded description of the derivative's ICC color space profile (ISO 15076-1), if supported by the format.

R16.4 All files SHOULD include geospatial information, if available and supported by the format.

4.1.2 Metadata

The following recommendations are specific to considerations of metadata.

R17 Data SHOULD be accompanied by internal and/or external descriptive metadata documenting aesthetic/engineering content and architectural/administrative context of the data; technical metadata documenting the formal characteristics of the data and its intemediating software; provenance metadata documenting the change history of the data over time; and intellectual property rights metadata documenting legal assertions associated with the data regarding copyright, contractual obligations, trade secrets, licenses, and terms of use.

R17.1 Descriptive metadata MAY be provided in terms of Describing Archives: A Content Standard (DACS) (SAA, 2019) or Dublin Core (ISO 15836-1).

R17.1.1 Descriptive metadata SHOULD conform to the general principles established by Cataloging Cultural Objects (Baca et al., 2006).

R17.1.2 Descriptive terms SHOULD conform to usage established by the Art and Architecture Thesaurus (AAT) and Cultural Objects Name Authority (CONA) (Getty Research Institute, 2021).

R17.2 Geospatial metadata MAY be provided in terms of the Content Standard for Digital Geospatial Metadata (FGDC, 1998).

R17.2.1 Geospatial place names SHOULD conform to usage established by the Thesaurus of Geographic Names (TGN) (Getty Research Institute, 2021).

R17.3 Technical metadata SHOULD be provided in terms appropriate for the underlying genre and format, for example, ISO 82045-5 for CAD (ISO, 2005); ANSI/NISO Z39.87/MIX for still images (ANSI/NISO, 2006; Library of Congress, 2008); the evolving Community Standards for 3D Data Preservation (as they are formalized) for 3D/AR/VR (CS3DP, 2021); DocumentMD for documents (Chou & Goethals, 2009); TextMD for text (Library of Congress, 2007); DataCite or DDI for data (DataCite, 2021; DDI, 2009); and the ISO 19115 family for geospatial information (ISO, 2014).

R17.4 Provenance metadata MAY be provided in terms of the general PROV model (Gill & Miles, 2013) or preservation-specific PREMIS data dictionary (OCLC/RLG, 2005)

R17.5 Rights metadata MAY be provided in terms of the Getty data dictionary (Whalen, 2016) or PREMIS (Coyle, 2006).

R17.6 Copyright status SHOULD be documented in terms of a Standardized Rights Statement (SRS) (Gore & Keller, 2018).

R18 Metadata SHOULD be created/acquired contemporaneously with the creation/acquisition of the described data.

R19 Metadata SHOULD be created/verified by the architectural or archival agent involved in the described data's creation/acquisition.

4.1.3 Packaging and Transfer

The following recommendations are specific to considerations of the packaging of architectural data for transfer into and out of archival systems and services.

R20 The package SHOULD conform to more specific recommendations of the Building for Tomorrow Transfer working group (*add citation here and the reference list below*).

R21 The package MUST incorporate or be accompanied by an inventory manifest that documents the complete list of all files, file sizes, and message digest algorithm types and values.

4.2 Software

The following recommendations are specific to considerations of intermediating software critical to the ability to open, render, and render preserved data.

R22 The architectural/engineering/facilities archival community SHOULD collectively engage with software vendors to encourage long-term third-party escrow of software, to be made available in the event that it is no longer commercially obtainable for permanent, worldwide, non-exclusive, no-cost, non-DRM- encumbered use by responsible archival programs for purposes of long-term preservation and eventual non-commercial scholarly (re)use of preserved design data.

R23 The architectural/engineering/facilities archival community SHOULD collectively engage with software vendors to obtain permanent, worldwide, non-exclusive, no-cost licenses for non-DRM-encumbered software for use by responsible archival programs for purposes of long-term preservation and eventual non-commercial scholarly (re)use of preserved design data.

R24 The architectural/engineering/facilities archival community SHOULD collectively encourage and support development and sustained operation of public online technical platforms for acquiring, preserving, and providing access to significant software

products. These platforms MUST operate in accordance to relevant copyright and contractual terms but SHOULD claim and exploit all legitimate preservation prerogatives under fair use doctrine.

R25 The architectural/engineering/facilities archival community SHOULD collectively encourage and support research, development, experimentation, and productionization of software emulation tools, systems, and services -- such as, but not restricted to, Emulation-as-a-Service- Infrastructure (EaaSI) (Anderson et al., 2018) -- as core strategic components of preservation activity.

R26 Archived software SHOULD be described in terms of metadata consistent with the evolving principles and standards defined by the Software Preservation Network's Metadata working group (SPN, 2021).

4.3 Legal

The following recommendations are specific to considerations of legal issues complicating or impeding preservation activity.

R27 The design and implementation of these recommendations MUST be legally defensible, in terms of copyright, contract, and fair use doctrine, balancing the legitimate intellectual property rights of data and software owners with the scholarly and public policy benefits of long-term stewardship of and access to architectural heritage.

R27 The AIA model contract (AIA, 2017) SHOULD be modified to include explicit provision granting to practitioners and/or clients a permanent, worldwide, non-exclusive, no-cost license to transmit curatorially-selected data to a responsible archival program for purposes of long-term preservation and eventual non-commercial scholarly (re)use, independent of other assertions or assignments of intellectual property rights ownership of that data.

R28 Archival donor agreements SHOULD be negotiated with explicit terms granting permanent permission to a responsible archival program to accept, reformat, copy, and distribute appropriate copies of curatorially-selected data for purposes of long-term preservation and eventual non-commercial scholarly (re)use, independent of other assertions or assignments of intellectual property rights ownership of that data.

R29.1 Donor agreements SHOULD conform to the recommendations of the SAA Digital Design Records Task Force (Leventhal et al., 2018).

R29.2 Donor agreements SHOULD enact the fullest acceptable transfer of intellectual property rights from the donor to the archival program.

R30 The architectural preservation community SHOULD actively promote and act according to software fair use principles for long-term preservation in scholarly contexts (ARL, 2019).

4.4 Discovery and Access

R31 Reference to archived data in discovery services MUST be accompanied by clear statements of legal ownership and controlling terms of use.

References

AIA B101 (2017), *Standard Form of Agreement Between Owner and Architect* <https://www.aiacontracts.org/contract-documents/25141-owner-architect-agreement>.

AIA C106 (2013), *Digital Data Licensing Agreement* <https://www.aiacontracts.org/contract-documents/18516-digital-data-licensing>.

AIA E203 (2013), *Building Information Modeling and Digital Data Exhibit* <https://www.aiacontracts.org/contract-documents/19026-building-information-modeling-and-digital-data-exhibit>.

Anderson, Seth; Cochrane, Euan; Thornton, KAtherine; Meyerson, Jessica; Rechert, Klaus; and Gate, Ethan (2018), *EaaSI (Emulation as a Service Infrastructure)* <https://osf.io/6mc4f/>.

ANSI/NISO Z39.87 (2006), *Data Dictionary - Technical Metadata for Digital Still Images*, <http://www.niso.org/publications/ansiniso-z3987-2006-r2017-data-dictionary-technical-metadata-digital-still-images>.

ARL (2019), *Code of Best Practices for Fair Use in Software Preservation*, <https://www.arl.org/wp-content/uploads/2018/09/2019.2.28-software-preservation-code-revised.pdf>.

AXELOS (2019), *ITIL Foundations*, 4th ed. <https://www.axelos.com/best-practice-solutions/itil>.

Baca, Murtha; Harpring, Patricia; Lanzi, Elisa; McRae, Linda; and Whiteside, Ann (2006), *Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images* (Chicago: ALA), <http://vraweb.org/wp-content/uploads/2020/04/CatalogingCulturalObjectsFullv2.pdf>.

Bradner, S. (1997), *Key Words for Use in RFCs to Indicate Requirement Levels*, RFC 2119, BCP 14 <https://www.ietf.org/rfc/rfc2119.txt>.

Chou, C.; and Goethals, A. (2009), *Document Metadata: Document Technical Metadata for Digital Preservation* <https://web.archive.org/web/20100508071650/http://fclaweb.fcla.edu/uploads/Lydia%20Motyka/FDA_documentation/documentMD.pdf>.

CIPA DC-008 (2019), *Exchangeable Image File Format for Digital Still Camera: Exif Version 2.32* <http://cipa.jp/std/documents/download_e.html?DC-008-Translation-2019-E>.Coyle, Karen (2006), *Rights in the PREMIS Data Model*, Library of Congress <https://www.loc.gov/standards/premis/Rights-in-the-PREMIS-Data-Model.pdf>.

CS3DP (2021), *Metadata*, Community Standards for 3D Data Preservation <https://cs3dp.org/metadata/>.

DataCite (2021), *DataCite Metadata S*chema Documentation for the Publication and Citation of Research Data and Other Research Outputs, version 4.4, <https://doi.org/10.14454/3w3z-sa82>.

DDI (2009), *DDI Lifecycle*, version 3.1 <https://ddialliance.org/Specification/>.

FCDC (1998), Content Standard for Digital Geospatial Metadata, version 2 <https://www.fgdc.gov/standards/projects/metadata/base-metadata/v2_0698.pdf>.

Getty Research Institute (2021), *Getty Vocabularies* <http://www.getty.edu/research/tools/vocabularies/index.html>.

Gill, Yolanda & Miles, Simon, eds. (2013), *PROV Model Primer*, W3C Working Group Note <http://www.w3.org/TR/prov-primer/>.

Gore, Emily & Keller, Paul, eds. (2018), *Recommendations for Standardized International Rights Statements* <https://rightsstatements.org/files/180531recommendations_for_standardized_international_rights_statements_v1.2.2.pdf>.

ISO 14721 (2012), *Space data and information transfer systems – Open archival information system (OAIS) – Reference model* <https://public.ccsds.org/Pubs/650x0m2.pdf>.

ISO 15076-1 (2010), *Image technology colour management — Architecture, profile format and data structure — Part 1: Based on ICC.1:2010* <http://www.color.org/specification/ICC1v43_2010-12.pdf>.

ISO 15489-1 (2016), *Information and documentation – Records Management: Part 1, Concepts and principles*.

ISO 15836-1 (2017), *Information and documentation — The Dublin Core metadata element set — Part 1: Core elements*.

ISO 16363 (2012), *Space data and information transfer systems – Audit and certification of trustworthy digital repositories* <https://public.ccsds.org/Pubs/652x0m1.pdf>.

ISO 16684-1 (2012), *Graphic technology – Extensible metadata platform (XMP) specification – Part 1: Data model, serialization and core properties* <https://www.adobe.com/devnet/xmp.html>.

ISO 19115-1 (2014), *Geographic information — Metadata — Part 1: Fundamentals* <https://www.iso.org/standard/53798.html>.

ISO 82045-5 (2005), *Document management — Part 5: Application of metadata for the construction and facility management sector* <https://www.iso.org/standard/34952.html>.

ISACA (2020), *COBIT* <https://www.isaca.org/resources/cobit>.

Lee, Kee Young; and Albert, Kendra (2021), *A Preservationist's Guide to the DCMA Exemption for Software Preservation*, Software Preservation Network/Cyberlaw Clinic @ Berkman Klein Center, Harvard University <https://www.softwarepreservationnetwork.org/a-preservationists-guide-to-the-dmca-exemption-for-software-preservation/>.

Leventhal, Aliza (2013), "Collaborative Preservation Efforts of CAD/BIM Files," *IS&T Archiving Conference*, Washington, April 2-5 (Springfield: IS&T), pp. 251-254 <https://www2.archivists.org/sites/all/files/IS&TPaper_Aliza%20Leventhal.pdf>.

Leventhal, Aliza; Schroffel, Laura; and Thompson, Jody (2018), "Deeds of Gift as a Tool to Faciliate Born Digital Design File Processing and Preservation," *Archive*Records 2018*, SAA Research Forum, Washington, August 14 <https://www2.archivists.org/sites/all/files/Leventhal_Schroffel_Thompson_Deeds%20FINAL.pdf>.

Library of Congress (2008), *MIX - NISO Metadata for Images in XML Schema*, version 2.0 <https://www.loc.gov/standards/mix/>.

Library of Congress (2007), *TextMD - Technical Metadata for Text*, version 2.2 <https://www.loc.gov/standards/textMD/>.

Library of Congress (2017), *Sustainability of Digital Formats: Planning for Library of Congress Collections* <https://www.loc.gov/preservation/digital/formats/index.shtml>.

Library of Congress (2018), *Designing the Future Landscape: Digital Architecture, Design & Engineering Assets*, Architecture, Design and Engineering Summit, Washington, November 16-17 <https://labs.loc.gov/static/labs/work/reports/Aleventhal-DesigningTheFutureLandscapeReport.pdf>

MIT (2009), *Final Report for the MIT FACADE Project, October 2006 - August 2009* <https://www.vai.be/volumes/general/FACADEFinalReport.pdf>.

NARA (2020), *Digital Preservation Framework for Risk Assessment and Preservation Planning* <https://www.archives.gov/preservation/electronic-records/digital-preservation-risk>.

OCLC/RLG (2005), *Data Dictionary for Preservation Metadata*, version 1.0 <https://www.loc.gov/standards/premis/v1/premis-dd_1.0_2005_May.pdf>.

SAA (2019), *Describing Archives: A Content Standard*, DACS 2019.0.3 <https://files.archivists.org/pubs/DACS_2019.0.3_Version.pdf>

SPN (2021), *Metadata: Enabling Discovery, Preservation, Interoperability, and Reuse of Software*, <https://www.softwarepreservationnetwork.org/core-activities/metadata/>.

Whalen, Maureen (2016), *Rights Metadata Made Simple*, Getty Research Institute <https://www.getty.edu/publications/intrometadata/rights-metadata/>.

Whiteside, Ann (2018), *Building for Tomorrow Forum Report* <https://projects.iq.harvard.edu/files/buildingtomorrow/files/building_for_tomorrow_-_forum_report.pdf>

Appendix 6: Transfer Guidelines and Resources

Introduction

Building for Tomorrow is a project led by the Frances Loeb Library at the Harvard University Graduate School of Design that aims to frame a collaborative infrastructure to support long-term preservation of digital design data. "Infrastructure" includes the technologies and tools required for digital archiving, preservation repositories, addressing intellectual property issues, and engaging the stakeholder communities – software vendors, designers, architects, architectural historians, archivists, librarians, technologists, and digital preservationists – in working towards shared understanding and agreement on how to best preserve and ensure accessibility to digital design records into the future. The project is supported by an IMLS National Forum Grant under the National Digital Platform funding priority.

Purpose and Scope

The purpose of the guidelines is to provide repositories with a set of extensible best practices and workflows for the transfer and ingest of born digital design records from creators to repositories. The guidelines acknowledge that not all repositories will have the same level of resources and capacity to acquire born digital material, and therefore are meant to support repositories at their point of need.

The guidelines build off of existing resources to examine the unique challenges posed by explicitly born digital design file types such as those generated by Computer-Aided Design (CAD) software. The guidelines outline and detail the resources, tools, and workflows needed to carry out the transfer and ingest of born digital material. The sections on conversations with the donor and repository resources are useful to all repositories doing this work, but are especially geared towards repositories in the initial planning phases. The technical considerations and accession sections identify and describe the roadblocks repositories may encounter when faced with digital design data and the workflows that could be used to overcome them.

The guidelines do not discuss the processing or description of born digital design records. For guidance on born digital processing and description work, see: the Society of American Archivist Design Records Section's Descriptive Elements for Born-Digital Records in Architectural Collections, UC Guidelines for Born-Digital Archival Description, and The Digital Processing Framework.

Terminology:

Computer-Aided Design (CAD) is a broad term that describes the use of computers to assist and aid in design processes. It can also refer, more specifically, to the software used for creating 2D drawings or 3D models of physical objects. The term CAD is sometimes colloquially used in architecture to designate 2D CAD drawings or describe all non-BIM architectural models.

Computer-Aided Architectural Design (CAAD) is a term that emerged in the 1980s to describe programs developed specifically for the architectural industry that have specialized architectural databases and support the creation of architectural objects.

Computer-aided manufacturing (CAM) is the use of software to control and manage the operations of manufacturing tools. It is often used to refer to software that translates CAD files into manufacturing instructions to automate numerically controlled (CNC) machine tools. A CAD/CAM software combines both computer-aided design and computer-aided manufacturing capabilities.

Building Information Modeling (BIM) is the process that utilizes a set of technologies, tools, and policies enabling multiple stakeholders to collaboratively design and manage a digital representation of an object or a project. It is important to note that BIM is not a software, a tool, or even a set of tools but rather a process that incorporates various tools. Moreover, a BIM model is a 3D representation of an object and its physical and functional characteristics. In other words, while a traditional CAD software deals with individual drawings or models, "BIM, by contrast, integrates all individual elements into an interconnected whole, and stores each element together with its metadata: information about its materiality, colour, manufacturer, or unit cost, for example. Hence, graphic (geometry) and non-graphic information (metadata) are linked together." (Ball 2013, 67) For example, while a traditional 3D model of a door, for example, is a three-dimensional representation of geometric form, a BIM model of a door incorporates the physical and functional characteristics of a door that can be represented in various forms. Additionally, BIM works on different levels (0-3) that reflect various degrees of complexity and collaboration that range from 2D CAD drafting with no collaboration between the stakeholders (level 0) to full collaboration between the stakeholders using a shared online project model (level 3).

Geographic Information Systems (GIS), as defined by The GIS Encyclopedia, is "An integrated collection of computer software and data used to view and manage information about geographic places, analyze spatial relationships, and model spatial processes. A GIS provides a framework for gathering and organizing spatial data and related information so that it can be displayed and analyzed."

Resources:

- Hovestadt, Ludger, Hirschberg, Urs and Fritz, Oliver. *Atlas of Digital Architecture*, Berlin, Boston: Birkhäuser, 2020.

- Ball, Alex. "Preserving Computer-Aided Design (CAD)" DPC *Technology Watch Report 13-02*. April 2013
- *BIM glossary of terms*, Designing Buildings Wiki (https://www.designingbuildings.co.uk/wiki/BIM_glossary_of_terms)
- GIS Glossary, The GIS Encyclopedia (http://wiki.gis.com/wiki/index.php/GIS_Glossary/G#GIS)
- The Ultimate Glossary of BIM Terms You Should Know, United BIM (https://www.united-bim.com/the-ultimate-glossary-of-bim-terms-you-should-know/)

Repository Resources

One of the most important parts of acquiring born-digital materials is first determining whether your repository is able to acquire these materials or if additional planning and resource allocation needs to happen first. This section will give a brief overview of things a repository should consider when thinking about accepting digital design files and will include links to resources that already exist or are forthcoming.

Labor

Before any of the sections below, it is important to consider the amount of labor involved in stewarding born-digital materials of any format. If a repository does not already have an existing program to acquire digital materials, the development of this infrastructure will take time depending on the resources and staff available. For repositories with an already established digital accessioning program, there can be other variables in place, as certain accessioning activities for born-digital material can take longer depending on how the repository is receiving digital files. For instance, ingesting files that are stored on physical carriers (optical disks, hard drives) can be a repetitive, time-consuming activity.

Reviewing and rethinking staffing models to ensure that a single individual is not responsible for the entirety of this labor can help prevent continued backlogs. It is also possible for some activities to be completed by library staff (students, assistants, etc.) so that an archivist can focus on other parts of processing.

Training & Professional Development

Depending on the kind of collecting institution, expertise and professional development needs will vary. For archivists responsible for the acquisition and accessioning of born-digital materials, training related to the tools and software used to do this, alongside experience with developing and maintaining workflows is necessary. However, for processing archivists training could just consist of learning how to arrange and describe born-digital records.

Since there are such a variety of institutions, collecting digital design records specifically – facilities, special collections and archives, museums – different skill sets and levels of expertise with born-digital materials will need to be cultivated. Continuing professional

development and training is also necessary for archivists doing this work, as designers explore and experiment with new technologies.

Resources:

- Descriptive Elements for Born-Digital Records in Architectural Collections
- Purcell, Aaron D. *The Digital Archives Handbook: A Guide to Creation, Management, and Preservation*. Lanham: Rowman & Littlefield, 2019.
- Ryan, Heather. *The No-Nonsense Guide to Born-Digital Content*. No-Nonsense Guides (Facet Publishing). London: Facet, 2018.
- Professional Organizations:
- Society of American Archivists - Design Records Section
- BitCurator Consortium (BCC)
- Digital Library Federation (DLF) Born-Digital Access Working Group

Storage & Cost

Just like collections of physical materials, planning for the storage and cost of maintaining born-digital archival materials should be accounted for early in the planning process.

Resources:

Total Cost of Stewardship: Responsible Collection Building in Archives and Special Collections
Pendergrass, Keith L., Walker Sampson, Tessa Walsh, and Laura Alagna. 2019. "Toward Environmentally Sustainable Digital Preservation." The American Archivist 82 (1): 165–206

Preservation and Access

As noted in other sections of the Building for Tomorrow paper, conversations with preservationists around the long-term preservation of digital design records is needed. While the bits of these files can be preserved, a greater understanding of the files is necessary in order to ensure that these records can be accessed in the future.

Resources:

Ball, Alex. "Preserving Computer-Aided Design (CAD)" DPC Technology Watch Report 13-02. April 2013

Emulation-as-a-Service Infrastructure

https://www.softwarepreservationnetwork.org/emulation-as-a-service-infrastructure/

Breitwieser, Kelly Stewart & Stefana. "SCOPE: A Digital Archives Access Interface." *The Code4Lib Journal*, no. 43 (February 14, 2019). https://journal.code4lib.org/articles/14283.

Knowledge Base

Due to the specialized technology and software used to create digital design files, many archivists might not have practical experience in handling some of these formats. While an archivist knows what a text or image file should look like and what software to use to make these files accessible, additional training on design specific software and tools may be necessary.

The Society of American Archivists has started to address the need for design records specific training with its course on *Managing Physical & Digital Architecture, Design, and Construction Records*. The SAA Design Records Section has also started to create resources that link archivists to tutorials related to design software and record creation.

Resources:

Managing Physical and Digital Architecture, Design, and Construction Records https://www2.archivists.org/prof-education/course-catalog/managing-physical-digital-architecture-design-and-construction-records
Archiving digital architectural records: national framework trial https://www.unisa.edu.au/connect/galleries-museums-and-centres/architecture-museum/archiving-digital-architectural-records-national-framework-trial/
Leventhal, A., Thompson, J., Anderson, A., Schubert, S., and Altenbach, A. (forthcoming) Design Records Appraisal Tool. *American Archivist*. Chicago: Society of American Archivists.

Conversations with the donor

In the pre-acquisition stage, it is critically important that repositories engage with donors to identify and discuss any born digital material for transfer. The proprietary nature of architecture and design software can present unique challenges to the acquiring repository, which is why these conversations are encouraged in the pre-acquisition stage.

Early donor or firm conversations allow repositories to gain a sense of the size and scope of the born digital material, and to assess whether or not they have the resources and capabilities to acquire born digital design records. If a repository does have the resources and capabilities, then these conversations will help them prepare for any technical requirements needed for the transfer of the material. When working with a firm, some of

this information may be gleaned through conversations with whomever is responsible for the maintenance of the firm's IT infrastructure.

When talking with the donor or firm, it's important for the repository to ask questions that will illuminate how the materials were generated, the software systems used to create the records, what file types are present, what file structures and naming conventions were used, if any files are password protected or contain PII, if there are any records not suitable for acquisition, and how the records have been stored and maintained. All of this information will help inform the repository on the best method of transfer and will also facilitate the accessioning process.

During the donor or firm conversation, it is also recommended to ask the donor or firm to supply any documentation of the records, such as: inventories, directory printouts, and/or supporting documentation related to the creation of the material, such as software user manuals. This will assist repositories in the post-acquisition stage, helping to verify that all records identified for transfer were indeed transferred.

Finally, the repository should work with the donor or firm to identify how the files will be prepared for transfer. In this step, the repository should determine any technical requirements and expertise the donor or firm may need in order to prepare the files and conduct their part of the transfer. For example, if a direct network transfer is identified as the preferred method of transfer, then the repository should make sure that the donor or firm is familiar with the software or tools being used to select and package the files.

Repositories may wish to conduct these conversations several ways, but it is recommended that they be recorded via a written survey, questionnaire, or report. Whenever possible, surveys and questionnaires should be sent to the donor or firm ahead of time, so that the donor or firm has ample time to prepare. Examples of donor surveys and questionnaires include the CCA's Submission of Digital Files Information Sheet (Appendix A) and the AIM's Donor Survey. In Appendix B,Below, see an example of how the AIM's Donor Survey can be adapted to born digital design records. Appendix C provides an example born-digital survey report that summarizes the findings of the in-person and written donor surveys, with a particular emphasis on how the findings will impact the transfer process.

Appendix 6.1: Adapted AIM's Donor Survey for Born Digital Design Records

1. General Work & Computing Habits

1.1. What are your chief activities? (e.g. writing, research, lecturing, architectural design)

1.2. What kinds of records do you create, maintain, and use in the course of each of these activities? (drafts of writings, research notes, lecture notes, correspondence, photographs, databases, digital modeling (CAD, BIM, other 3D objects), etc.)

1.3. Can you describe your general work habits with computers in support of these activities? (e.g. you write or draft first by hand, then input work into computer, you use different computers for different kinds of work, you're always online, etc.).

2. Digital Material Creation

2.1. Are you solely responsible for creating your digital files?

2.2. If not, who else is involved, and what are their roles?

2.3. Do you maintain digital files created by others? If yes, how do you separate your files and files created by others?

2.4. Do you share your computer with other people? If yes, how are files created by different people separated?

2.5. Do you separate your personal files from your work files? (eg: lecture files, architectural practice files, client files, etc).

2.6. What are the earliest and latest creation dates (roughly) of your digital files?

2.7. What kind of software environments do you use to create digital design files? (eg: AutoCAD, Microstation, BIM, etc.)

3. Varieties of Digital Material

3.1. What types of digital files are created? (e.g. word processing files, images, spreadsheets, databases, specific design files such as .dgn, or .dwg)

3.2. If you create files in both digital and paper formats, do certain files exist in both formats? (e.g. drafts of writings, email, etc.)

4. Digital Material Organization

4.1. How are digital files named?

4.2. Is some kind of version control used? (e.g. filename1, filename2, to represent 1st and 2nd drafts of the file.)

4.3. How are your digital files currently organized? (e.g. filed in named folders? by projects? by topics? some other scheme?)

4.4. Have you always had this organization? If not, can you summarize/characterize previous organizations, and roughly when and why you made changes?

4.5. Are digital files destroyed in regular intervals?

4.6. Do you use more than one computer (e.g. office desktop, office portable computer, home desktop, etc.)? If yes, how do you synchronize files between different computers?

5. Mobile Device

5.1. Do you use smart phones (e.g. Blackberries, iPhone, Android phone, etc.)? If yes, do you store contents in the smart phone elsewhere?

5.2. Do you use tablets PC (e.g. iPad, etc.)? If yes, do you store contents in the tablet PC elsewhere?

6. Email

6.1. Do you have multiple email accounts?

6.2. Which email programs/services are you using? (e.g. Email program provided by your work place, Outlook, Mac Mail, Hotmail, Gmail, Yahoo! Mail, etc.)

6.3. How is your email currently organized? (e.g. in self-created email folders, etc.)

6.4. Have you always had this organization? Do you use the sorting function with any regularity to re-order your email?

6.5. How is email saved? (e.g. untouched in the email program, a copy in your PC, printed out in paper, etc.)

6.6. Are email and paper correspondence managed together or separately?

6.7. Do you use address books?

6.8. Is there a space quota assigned to your email account? If yes, have you ever exceeded the quota assigned?

7. Webpages / Blogs – Stalled! webpage

7.1. How often is the Stalled! webpage updated?

7.3. What software do you use to update the Stalled! webpage?

7.4. Have copies (digital or paper) of previous versions been kept?

8. Photo / Video Sharing Sites (e.g. Flickr, Picasa, YouTube, etc.)

8.1. Do you post photos / videos to these web sites? If yes, which one?

8.2. How often do you post contents?

8.3. Do you delete photos / videos posted? If yes, do you have a copy of the deleted postings?

9. Document Sharing Sites (e.g. SlideShare, Scribd, Google Doc, etc.)

9.1. Do you post documents to these web sites? If yes, which one?

9.2. How often do you post contents?

9.3. Do you delete documents posted? If yes, do you have a copy of the deleted postings?

10. Digital Files Storage / Backup

10.1. Do you / your institution have a backup routine for your files / emails? If you don't know, do you mind we ask your technical support? How can we contact your technical support?

10.2. What media are used for backup files? (e.g. optical disk, hard disk, file server, web based backup service such as SugarSync., etc.)

10.3. Do you transfer files in your old computer to your new computer? If yes, what types of files are transferred? Did you encounter any problems in transferring the files?

10.4. Do you keep your old computers? Roughly when were they being used? Can you tell us what platforms they run on?

10.5. Have you ever experienced a serious hardware failure (e.g. hard-drive crash)? If yes, are the files in the affected computer recovered?

10.6. Are any digital files stored in unusual storage media? (e.g. punch cards, 8 inch. floppy diskettes, etc.)

11. Privacy and security

11.1. Are some digital file types of a sensitive nature? (e.g. tax records, medical records, peer-review comments, letters of recommendation, student records, etc.)

11.2. Are there files that you would want destroyed? If yes, please provide details so that we can act upon when we encounter 3such files when processing your files.

11.3. Do any digital files require passwords?

11.4. Where are user names and passwords kept? What service / software are used to save them?

11.5. Do you use digital watermarks? On what types of digital files? For what reasons?

12. File Transfer Arrangement

12.1. Do you want to delete any files / re-organize the files before the transfer?

12.2. Are there files you would like to transfer to us later? When?

Appendix 6.2: Sample Born Digital Survey Report

This is an anonymized report – any identifying information such as firm and donor names, archival repository names, and project names, have been removed or changed for privacy. Therefore:

- The Firm: refers to an architectural design practice.
- The Donor: refers to the founder of The Firm and the signatory of the Deed of Gift.
- Archives Staff: refers to the professional archivists working at an archival repository.

Sections:

1. Objective
2. Summary
3. Records
    a. Project drive
    b. Marketing drive
    c. Associated material on The Donor's personal drive
    d. The Firm's Website
    e. The Donor's email
    f. Born digital materials stored on CD
4. Next Steps

1. Objective:

To identify all electronic records on The Firm's server relating to the projects, research, and writings previously selected by Archives Staff for acquisition; to gain a detailed understanding of the existing file structure, file formats, and naming conventions; and to identify any potential technical issues related to transfer, preservation, and access.

2. Summary:

On October 1st, 2018, Archives Staff visited The Firm's offices in Massachusetts. Archives Staff met with The Firm's Assistant and Senior Associate to survey born digital records for the following areas: research files, project files, organizational records, and The Firm's website. Archives Staff discussed, but did not view, The Donor's email account.

This meeting follows a site visit Archives Staff conducted at The Firm's offices to survey analog archival material. At that visit Archives Staff learned that The Firm's research, writing, and project records exist in three states: analog only, electronic only, and both analog and electronic. Archives Staff determined that an additional visit dedicated to surveying the born digital records was necessary.

With the AIMS Born Digital Survey as a guide, Archives Staff were able to accomplish the objectives stated above in the meeting. The records identified for acquisition exist in six areas:

1. Project drive
2. Marketing drive
3. Associated material on The Donor's personal drive
4. The Firm's website
5. The Donor's email
6. Pre-2008 born digital materials stored on CD

These six categories will be used to describe the types of records identified for acquisition as well as the technical aspects of transfer, preservation, and access. Each category is distinct from the others, presenting its own set of issues related to records creation, maintenance, and preservation. However, there were two overall observations made that the categories have in common.

- Records destruction: The Donor typically does not delete anything. Records that would otherwise be deleted are instead stored in folders labeled "superseded."

- The final copies of all associated materials for project records, writings, and research are kept on the server.

The final section of this report will describe what Archives Staff perceive the next steps to be, which includes the technical aspects of transfer, as well as the creation of a Born Digital Addendum to the Deed of Gift.

3. Records

3.a. Project drive

Records for acquisition: All project records related to the following projects:

- City Library
- City Loft
- University Field House
- Private Residence

The project files contain 2D and 3D models, design development drawings, construction documentation, correspondence (client and consultant, in-bound and out-bound), meeting notes, schedules, images, site photos, specifications, bid documents, and final as-built drawing sets. Each phase of the project is documented.

File Structure: The drive contains top level folders for each project, labeled by project name and number. Within each project folder are sub-folders that identify each phase of the project, and within those sub-folders are sub-sub-folders containing drawings, correspondence, photographs, and construction documents. Drawing sets that have been superseded are labeled as such at the folder level. The Project drive structure is consistent across all project folders.

File Formats: The Firm uses Rhino to produce 3D modeling design files and AutoCad to produce 2D drawings. The Firm staff has ensured that final drawing sets created in AutoCad are saved so that any external references to images, fonts, and other external libraries are bound in one single file. In addition to this final AutoCad file, The Firm staff save a final 3D PDF file, which is exported from Rhino. Other file formats in these folders include Word documents, Powerpoint, and PDF.

Naming Convention: Each version of a design created in any software environment is labeled by date of creation and author. This naming convention is consistent across all projects.

Access: Each project folder has an Admin Folder, which contains administrative information on the building such as codes, client information, contacts, meeting notes, and project schedule. Due to the amount of Personal Identifying Information (PII) on clients, such as addresses and phone numbers, Archives Staff suggest that all Admin files be restricted for a time period to be determined. Client information does not exist in other project records.

Future researcher access to the Rhino and AutoCad files can be addressed in two ways. The first is to offer immediate access to the 3D PDF files, which only requires the user to view the files in Adobe. The second is to offer emulation as a service to the actual files in their native software environment(s). This option will require purchasing a license to Rhino.

Transfer: The entire size of the Project drive is less than 2TB. A proposed physical transfer method could be an onsite transfer of the digital files from The Firm onsite servers directly to a hard drive (provided by Archives Staff) using either logical forensic imaging methods or simple file transfer with integrity checking.

3.b. Marketing drive

Records for acquisition: The presentation and marketing files (ex: Requests for Information and presentation drawings) related to the project records from section 3.a. are stored on this drive (as opposed to the Project drive) and should be selected for acquisition. In addition to these records, there are also more routine administrative records which Archives Staff would not want to keep. Separating the records related to project records from the administrative records would be difficult due to the file structure. Archives Staff suggest identifying the folders related to the projects listed in section 3.a. for immediate acquisition and conducting further appraisal after transfer.

File Structure: The file structure on this drive is not standard across folders and sub-folders. As opposed to the Project drive, files within a project folder have no clear arrangement.

File Formats: Administrator, Photoshop, Word, and Powerpoint are the main file formats in this drive. For Powerpoint presentations, external references are embedded into the final file (example: a PowerPoint that uses animation or videos).

Naming Convention: There is no standard naming convention used in this drive, however most files appear to be clearly labeled, which makes identification of what each individual folder and file is a little easier.

Transfer: A proposed physical transfer method could be an onsite transfer of the digital files from The Firm's onsite servers directly to a hard drive (provided by Archives Staff) using either logical forensic imaging methods or simple file transfer with integrity checking.

3.c. Associated material on The Donor's drive

Records for acquisition: On The Donor's drive are the final versions of research, writings, lectures, presentations, and workshops. Archives Staff have identified lecture and workshop folders for acquisition, with further appraisal needed after acquisition.

File Structure: The general structure on this drive is: date, name of lecture/workshop.

File Formats: Word, Powerpoint, audio/visual formats such as .mp3 and .wav

Naming Convention: Final lectures are named similarly to the folder: date, name of lecture/workshop. However, there is no clear naming convention across this drive.

Access: Recordings of lectures and workshops may have rights issues for re-publication online if other participants were recorded without license.

Transfer: A proposed physical transfer method could be an onsite transfer of the digital files from The Donor's onsite servers directly to a hard drive (provided by Archives Staff) using either logical forensic imaging methods or simple file transfer with integrity checking.

3.d. The Firm's Website

Records for acquisition: The Firm's website is used to promote The Donor's lectures, workshops, writings, interviews, and research. The website recently launched, with updates made on a quarterly basis. Archives Staff recommend web archiving the published public version, pay for historic crawls of the url, and then setup bi-annual or quarterly crawls into the future.

3.e. The Donor's Email

Records for acquisition: The Donor's personal and professional email is held on one Gmail account. The Donor does not delete any emails and utilizes the "Archive" feature when a project is completed. The result is an inbox of approximately 10-20k emails, with both personal and professional correspondence mixed in. If The Donor or The Firm's staff need to search for a particular email, they search by correspondent name. Archives Staff were

unable to confirm if the Gmail account captures both ingoing/outgoing correspondence, or just outgoing.

Archives Staff were unable to confirm if there is any overlap between the Gmail account and correspondence on the Project drive. If the Gmail account is acquired, it would largely be a fishing expedition for relevant correspondence, and that work would fall on Archives Staff. This scenario could help Archives Staff test various email tools and be used as a practical use-case in email acquisition, transfer, arrangement and description. Archives Staff need to determine the resource capacities (mainly staff time and expertise) to undertake this work are available.

File Structure: Emails are either in the Inbox, or in the "Archive" folder. There is no hierarchical structure to the account.

Naming Convention: No tags are used on emails.

Transfer: One transfer method would be to have The Firm staff save extracted email to a cloud storage space that could be shared with Archives staff (ie: Google Drive). The email would then be ingested to a preservation system, which could present some technical issues.

3.f. Born digital materials stored on CDs with analog collection material

Records for acquisition: These CDs are physically located with the analog archive. They contain project records, but Archives Staff do not know much else about the contents. The Firm staff indicated that for most projects, the final draft should have been migrated and saved to the The Firm's server, but any files leading up to the final draft (design sketches, drafts, correspondence, etc) would be on the CD. There is one box of CDs.

File Structure: File structure for the CDs is not known.

File Format: The Firm's staff said earlier digital modeling software, such as Sketchup and FormZ, would be found on these CDs.

Naming Convention: Naming convention is not known.

Access: The major concern for access is the presence of Sketchup and FormZ. Both software environments are highly proprietary and can only open files in one of the most current versions – meaning that it might not be possible to open these files at all, even though both software companies still exist and are active on the market. Archives Staff do not know if emulation as a service will be able to provide access to these files.

Transfer: The proposed transfer method is the same as the paper file transfer with forensic disk imaging following current Archives Staff practice.

4. Next Steps:

The next step is to follow up with The Donor regarding any overlap between the Gmail account and the Project drive. This should inform how Archives Staff move forward with the Gmail account.

Archives Staff will also need to draft an Addendum to the Deed of Gift to outline the records for accessioning and the transfer procedures. The Firm's staff have provided file directory printouts for the Project and Marketing drives to help in the selection process.

Appendix 6.3: Transfer Resources

Pre-Acquisition Work

Conversations with the donor or firm before acquiring material help to identify key components of the transfer process. Having a survey or questionnaire completed by the donor/firm gives the donors and collecting institutions important information to guide future decisions regarding the transfer of material. What works for one donor/firm may not work for another, so it is useful to have workflows that address a variety of possibilities. In order to develop a transfer plan, determine with the donor the following:

- **Current Storage Location of the Donor's files** - Born-digital files can be stored in several different types of digital storage largely dependent on when they were created, what the donors digital storage infrastructure looks like, and how recently they were considered to be active files that may need to be referenced frequently. Files created before the prevalence of affordable local servers and/or cloud-based storage are likely to be stored on physical carriers such as floppy disks, optical media, USBs, and hard-drives. While there are a lot of resources documenting the transfer of files from these carriers during ingest, transferring files from a private donor's local server or cloud-based storage system is a less documented activity subject to many variables.

- **Method of Transfer** - Establishing how the files will be transferred once any pre-acquisition appraisal has occurred and the donor has prepared the files will help determine what additional planning needs to occur. If the files are already on physical carriers, the donor/firm can just give this media to the collecting institution along with any supporting inventories or related documentation. Since donors do not always have the means to access this material it is particularly important for there to be conversations around Personally Identifiable Information (PII) when the donor is not entirely certain of the content of the carriers.

  For files that are stored on local storage or in a cloud-based storage system, there are several methods of transferring this material to the collecting institution. Depending on the amount of time and resources available to the donors they might choose to transfer the material by uploading it to a file hosting service, provide access to the collecting institution to their cloud-based file storage, copy items onto a hard-drive, or most often in cases of institutional records utilize local SFTP/FTP options. When deciding between these options, knowing the size of the accession is key as some options may have size restrictions.

If a donor will be spending some time organizing and preparing files for transfer they should be provided with additional guidance on how to do this quickly and efficiently while also minimizing the impact on original metadata as much as possible.

If using tools that need to be installed on a donor's local machine try and get this done well before the scheduled transfer date.

Tools and Applications for Transfers

Deciding which tools to use can depend on multiple factors such as a collecting institution's needs and goals, time and resources available to both the donor and collecting institution, and the anticipated size of the deposit. This overview is representative of some of the most common tools used for this activity and which have been used by the Building for Tomorrow team. These tools focus primarily on the transfer of files that are not stored on a physical carrier.

Packaging Applications

- Bagger - https://github.com/LibraryOfCongress/bagger
- Exactly - https://www.weareavp.com/products/exactly/

These applications both use the BagIt File Packaging specification (https://datatracker.ietf.org/doc/html/draft-kunze-bagit-14)

Disk Imaging Tools

Disk Imaging tools like FTK Imager can be installed and run on an external drive that is then connected to a donor's machine. This allows the archivist to create a disk image that can then be ingested into the repository's storage.

- FTK Imager
  https://support.accessdata.com/hc/en-us/articles/204275735-FTK-Imager-version-3-3-0-User-Guide
- Run FTK Imager from a Flash Drive
  https://support.accessdata.com/hc/en-us/articles/203681809-Run-FTK-Imager-from-a-flash-drive-Imager-Lite-

Copying Applications

Encouraging donors to install a copying application like Teracopy to ensure identical copies of files are being made as they prepare files for transfer.

- Teracopy
  https://www.codesector.com/teracopy

Sample Acquisition Workflows

The following resources contain detailed guidance and recommendations for digital preservation workflows including detailed sections on transfer tools and methods.

- Guidance for Digital Preservation Workflows https://www.nationalarchives.gov.uk/archives-sector/projects-and-programmes/plugged-in-powered-up/digital-preservation-workflows/
- Acquisition and Appraisal - Digital Preservation Handbook https://www.dpconline.org/handbook/organisational-activities/acquisition-and-appraisal
- OSSArcFlow: Guide to Documenting Born-Digital Archival Workflows https://educopia.org/wp-content/uploads/2020/06/OSSArcFlow_Guide_FINAL.pdf
- Walk This Way: Detailed Steps for Transferring Born-Digital Content from Media You Can Read In-house - OCLC https://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf
- How to Access Digital Files from the Nineties https://www.cca.qc.ca/en/articles/issues/3/technology-sometimes-falls-short/49023/how-to-access-digital-files-from-the-nineties

Future recommendations for work related to acquisition

- Draft Transfer Checklist

  In order to make sure that any tasks specific to transferring digital design files are completed during the acquisition process, one of the recommendations to come out of the November 2020 meeting was for the development of a Transfer Checklist that can be tailored to a collecting institutions needs and goals while being mindful of the labor required on the part of the donor or firm.

- Investigate how transfer tools already used by creators can be utilized for the transfer process.

  In order to simplify the transfer process, it might be worthwhile to look a closer at tools already being used in practice to transfer material. Explore whether or not it would be possible to incorporate these applications into the transfer process.

- Collect Use Cases and Workflows

Resources Related to Technical Considerations

1. Software and File Format Resources

   - A list of CAD file Format, Wikipedia (https://en.wikipedia.org/wiki/List_of_file_formats#Computer-aided_design)
   - CAD File format, Design Presentation Associates
   - File Formats for BIM, BIM Wiki
   - DWG family preservation and sustainability factors, Library of Congress

- Leventhal, Aliza "Architectural and Design Collections" in Aaron D. Purcell, ed., *The Digital Archives Handbook: A Guide to Creation, Management, and Preservation* (Lanham, Maryland: Rowman & Littlefield, 2019).
- Beyond TIFF and JPEG2000: PDF/A as an OAIS submission information package container

2. Best practices

    - LOC Recommended Formats Statement 2020-2021 https://blogs.loc.gov/thesignal/2014/08/untangling-the-knot-of-cad-preservation/

    - Untangling the Knot of CAD Preservation: https://blogs.loc.gov/thesignal/2014/08/untangling-the-knot-of-cad-preservation/

3. Export and Access Format

    In addition to the original CAD file, CAD drawings can be exported and preserved in other formats such as PDF and JPEG. The exact steps needed to export the drawings depend on the software. Following are resources on how to export CAD drawings created by some of the most widely used software programs.

    AutoCAD
    - To Export Selected Layouts to a PDF File https://knowledge.autodesk.com/support/autocad/learn-explore/caas/CloudHelp/cloudhelp/2020/ENU/AutoCAD-Core/files/GUID-5E91B638-7C4A-4CF3-8F7C-593368D72A97-htm.html

    Revit

    - Print to PDF https://knowledge.autodesk.com/support/revit-products/learn-explore/caas/CloudHelp/cloudhelp/2020/ENU/Revit-DocumentPresent/files/GUID-8B7424DD-C07A-4FD7-B4DB-5F7F6F14D8E8-htm.html

    Rhino
    - Rhino Export to other 3D File Formats http://docs.mcneel.com/rhino/5/help/en-us/fileio/_index_of_import_export_file_types.htm
    - Rhino to PDF https://wiki.mcneel.com/rhino/creatingpdfs
    - Rhino to JPG, PXC, PNG, BMP, TIFF, TGA http://docs.mcneel.com/rhino/5/help/en-us/commands/viewcapture.htm

Sketchup

- Exporting or Printing Your LayOut Document https://help.sketchup.com/en/layout/exporting-or-printing-your-layout-document

4. Embedded Information

In some cases, externally referenced drawing files (X-references), blocks, raster images are embedded in CAD files. Ensure that all externally referenced files are included in the package you are receiving or are bound to their respective drawings.

Xref in CAD files

- To Bind an Xref to the Current Drawing https://knowledge.autodesk.com/support/autocad/learn-explore/caas/CloudHelp/cloudhelp/2020/ENU/AutoCAD-Core/files/GUID-5BF89799-D496-4AAD-8773-9D4BF9B9

Image File in CAD files

- Embed an Image File into AutoCAD https://lynn.blogs.com/lynn_allens_blog/2015/07/embed-an-image-file-into-autocad.html

Custom Fonts

- Adding custom fonts in AutoCAD and transferring them. https://www.thesourcecad.com/add-font-autocad/#:~:text=Go%20to%20%E2%80%9CC%3A%5CProgram,the%20list%20of%20AutoCAD%20fonts
- Text missing or does not display in AutoCAD drawings | AutoCAD https://knowledge.autodesk.com/support/autocad/troubleshooting/caas/sfdcarticles/sfdcarticles/Text-disappeared-when-opening-drawings.html

Plug-ins

A plug-in is a software module that extends the functionality of CAD software. When plug-ins are used, ensure that all required files are included: the original CAD file and the plug-in file if needed. For example, Grasshopper, a visual programming language and environment that runs within the Rhinoceros 3D, requires a separate file that is opened/saved from within the Rhinoceros 3D interface. Within the Grasshopper environment, many other plug-ins are usually used that are necessary for opening the Rhinoceros 3D models they were used in producing. These plug-ins should be saved along

with the Rhino file. Package Restore in Grasshopper can be used to download and install missing plug-ins.

- Package Restore in Grasshopper https://developer.rhino3d.com/guides/yak/package-restore-in-grasshopper/

Many architects also preserve screenshots of the Grasshopper interface to document the visual algorithms they used in the design.

Appendix 7: Standards Resources

This list includes some of the most important standards of interest to the various parties involved in the archiving process of born-digital design data.

This is available online at the Building for Tomorrow website.

Reference	Title	Source	Year	Implementer / Primary Audience	Categories	Simplified Scope	Scope and Notes	Normative Reference (Indispensable Reference documents)	URL
ISO 14721:2012	Space data and information transfer systems — Open archival information system (OAIS) — Reference model	The International Organization for Standardization (ISO)	2012	Custodial Institution	Digital Repositories	Reference model for an open archival information system (OAIS).	OAIS is an archive, consisting of an organization, which may be part of a larger organization, of people and systems that has accepted the responsibility to preserve information and make it available for a designated community. It meets a set of such responsibilities as defined in this International Standard, and this allows an OAIS archive to be distinguished from other uses of the term "archive". The term "open" in OAIS is used to imply that ISO 14721:2012, as well as future related International Standards, are developed in open forums, and it does not imply that access to the archive is unrestricted.		https://www.iso.org/standard/57284.html
ISO 16363:2012	Space data and information transfer systems — Audit and certification of trustworthy digital	The International Organization for Standardization (ISO)	2012	Custodial Institution	Digital Repositories	Standard for Trusted Digital Repositories	Recommended practice for assessing the trustworthiness of digital repositories. It is applicable to the entire range of digital repositories. ISO 16363:2012 can be used as a basis for certification.		https://www.iso.org/standard/56510.html
PD 19650-0:2019	Transition guidance to BS EN ISO 19650	British Standards Institute (BSI)	2020	Information producers	Data Management and Sharing	Guidance for transitioning from the 1192 series to the ISO 19650 series	Provides information for those familiar with BS 1192:2007+A2:2016 and PAS 1192-2:2013 during the transition to BS EN ISO 19650-1 and BS EN ISO 19650-2. Covers useful terms and definitions, the role, structure and status, BIM level 2 and information management maturity, mapping, language used in ISO 19650 and how it relates in the UK	BS EN ISO 19650 1:2018, BS EN ISO 19650 2:2018	https://shop.bsigroup.com/ProductDetail/?pid=0000000003041181 7
BS EN ISO 19650-1:2018	Organization and digitization of information about buildings and civil	The International Organization for Standardization	2018	Information producers	Data Management and Sharing	Principles for information management using (BIM)	Recommendations for a framework to manage information including exchanging, recording, versioning and organizing for all actors.	None	https://www.iso.org/standard/68078.html

	engineering works, including building information modelling (BIM). Information management using building information modelling. Concepts and principles	(ISO)							
ISO 19650-2:2018	Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM). Information management using building information modelling. Delivery phase of the assets	The International Organization for Standardization (ISO)	2018	Information producers	Data Management and Sharing	Specification for information management during the delivery phase of a construction project using (BIM)	Specifies requirements for information management, in the form of a management process, within the context of the delivery phase of assets and the exchanges of information within it, using building information modelling.	ISO 19650-1, ISO 12006-2	https://www.iso.org/standard/68080.html
BS 8541-6:2015	Library objects for architecture, engineering and construction. Product and facility declarations. Code of practice	British Standards Institution (BSI)	2015	Information producers	Data Management and Sharing	Code of practice for the transmitting of product declarations within library objects	Relate to transmission of product and facility declarations of the construction library objects which will be used in the project and life cycle stages, mainly comprising manufacturer specific products	BS 8541-1, BS 8541-2, BS 8541-3, BS 8541-4, BS ISO 16739	https://shop.bsigroup.com/en/ProductDetail/?pid=000000000030294760&_ga=2.195653044.1781712144.1613055923-936222298.1613055923
ISO 29481-2:2012	Building information models. Information delivery manual. Interaction framework	The International Organization for Standardization (ISO)	2012	Information producers	Data Management and Sharing	Specification for an information delivery manual	is intended to facilitate interoperability between software applications used in the construction process, to promote digital collaboration between actors in the building construction process, and to provide a basis for accurate, reliable, repeatable, and high-quality information exchange.	ISO 29481-1	https://www.iso.org/standard/55691.html
ISO 29481-1:2016	Building information models. Information delivery manual.	The International Organization for Standardi	2016	Information producers	Data Management and Sharing	Principles for interoperability between software applications using information delivery manuals.	It is is intended to facilitate interoperability between software applications used during all stages of the life cycle of construction works, including briefing, design, documentation,	ISO 6707-1	https://www.iso.org/standard/60553.html

	Methodology and format	zation (ISO)					construction, operation and maintenance, and demolition. It promotes digital collaboration between actors in the construction process and provides a basis for accurate, reliable, repeatable and high-quality information exchange.		
PAS 1192-6:2018	Specification for collaborative sharing and use of structured Health and Safety information using BIM	The International Organization for Standardization (ISO)	2018	Information producers	Data Management and Sharing	Specification for capturing and sharing information relating to health and safety risks using BIM	It provides guidance on how health and safety information is produced, flows and can be used throughout the project and asset lifecycle. While all health and safety risk information can be included within an information model, this PAS requires the contextualization and filtering of hazards and risks to prioritize the elevated risks and aspects that are safety critical.	PAS 1192-2, BS 1192-4, PAS 1192-5, BS 8536-1, BS 8536-2	https://shop.bsigroup.com/ProductDetail?pid=000000000030347710
ISO 19650-3:2020	Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information modelling — Part 3: Operational phase of the assets	The International Organization for Standardization (ISO)	2020	Information producers	Data Management and Sharing	Specification for securing the sharing and delivery of project and asset information when using BIM	specifies requirements for information management, in the form of a management process, within the context of the operational phase of assets and the exchanges of information within it, using building information modelling.	BS ISO 55000, PAS 1192-2:2013, PAS 1192-3:2014, BS 1192-4:2014	https://www.iso.org/standard/75109.html
ISO 11091:1994	Construction drawings. Landscape drawing practice	The International Organization for Standardization (ISO)	1994	Information producers	Representation conventions	Specification for symbols and conventions for landscape drawings		None	https://www.iso.org/standard/19080.html
ISO 6284:1996	Construction drawings. Indication of limit deviations	The International Organization for Standardization (ISO)	1996	Information producers	Representation conventions	Specification for representing limit deviations on construction drawings		ISO 286-1:1988, ISO 1803, ISO 4068:1978, ISO 9431:1990	https://www.iso.org/standard/12568.html
ISO 3766:2003	Construction drawings. Simplified representation of concrete reinforcement	The International Organization for Standardization (ISO)	2003	Information producers	Representation conventions	Specification for representing concrete reinforcement on construction drawings		ISO 128-23:1999, ISO 10209-4:1999	https://www.iso.org/standard/34171.html
ISO 7518:1983	Construction drawings. Simplified representation of demolition	British Standards Institute (BSI)	1983	Information producers	Representation conventions	Specification for the representation of existing, proposed and demolished elements on		ISO 128, ISO 3098-1, ISO 4069	https://shop.bsigroup.com/ProductDetail?pid

	and rebuilding					construction drawings.			=000000000019983603
ISO 128-23:1999	Technical drawings. General principles of presentation. Lines on construction drawings	The International Organization for Standardization (ISO)	1999	Information producers	Representation conventions	Specification for the types of lines to be used on construction drawings.		ISO 128-20:1996, ISO 29:1985, ISO 3766:1995, ISO 4068:1978, ISO 4069:1977, ISO 4463-1:1989, ISO 6428:1982, ISO 7519:1991, ISO 8048:1984, ISO 8560:1986, ISO 11091:1994	https://www.iso.org/standard/22292.html
ISO 8560:2019	Technical drawings -- Construction drawings -- Representation of modular sizes, lines and grids	The International Organization for Standardization (ISO)	2019	Information producers	Representation conventions	Specification for the types of lines and their use for representing modular sizes	Allows information relating to modular sizes on construction drawings to be conveyed in a consistent manner	ISO 128-1, ISO 129-1	https://www.iso.org/obp/ui/#iso:std:iso:8560:ed-2:v1:en
BS 8541-2:2011	Library objects for architecture, engineering and construction. Recommended 2D symbols of building elements for use in building information modelling	British Standards Institute (BSI)	2011	Information producers	Representation conventions	Code of practice for symbols and conventions to use when producing construction drawings		BS 3975-4, BS EN ISO 9431	https://shop.bsigroup.com/ProductDetail/?pid=000000000030255370
BS 8541-3:2012	Library objects for architecture, engineering and construction. Shape and measurement. Code of practice	British Standards Institute (BSI)	2012	Information producers	Representation conventions	Code of practice for characterizing the shape and measurement of library objects	Provides recommendations for shape and measurement of construction library objects used in building construction and facility domains, setting levels of detail and representations of measurement.	BS 8541-1:2012	https://shop.bsigroup.com/ProductDetail/?pid=000000000030240258
BS 8536-1:2015	Briefing for design and construction. Code of practice for facilities management (Buildings infrastructure)	British Standards Institution (BSI)	2015	Information producers	Management of Construction Processes	Code of practice for briefing; taking into account soft landing and BIM principles	Part of the UK BIM Framework suite of documents developed to help the construction industry adopt BIM. It gives recommendations for the use of data and information needed for briefing for design and construction, ensuring that the future performance and use of a building is considered. The standard applies to all new buildings projects and major refurbishments.	BS 1192:2007, BS 1192-4, PAS 1192-2, PAS 1192-3, PAS 1192-5	https://shop.bsigroup.com/ProductDetail/?pid=000000000030315621
BS 8536-2:2015	Briefing for design and construction. Code of practice for asset	British Standards Institution (BSI)	2015	Information producers	Management of Construction Processes	Code of practice for briefing; taking into account soft landing and BIM principles	Part of the UK BIM Framework suite of documents developed to help the construction industry adopt BIM. It gives recommendations for the use	None	https://shop.bsigroup.com/ProductDetail/?pid

	management (Linear and geographical infrastructure)						of data and information needed for briefing for design and construction, ensuring that the future performance and use of a building is considered. The standard applies to all new buildings projects and major refurbishments.		=000000000030315621
BS 7000-4:2013	Design management systems. Guide to managing design in construction	British Standards Institute (BSI)	2013	Information producers	Management of Construction Processes	Guidance on management of the construction design process	It provides guidance on resources, record management, production and intellectual property when undertaking design management	BS 1192, BS 4778-2, BS 7000-10, BS EN 12973, BS EN ISO 10012, BS EN ISO 12006-2, PAS 1192-2	https://shop.bsigroup.com/ProductDetail?pid=000000000030314817
ISO 22263:2008	Organization of information about construction works. Framework for management of project information	The International Organization for Standardization (ISO)	2008	Information producers	Management of Construction Processes	Specification for organizing project information by outlining the key processes, sub-processes and activities undertaken	It specifies a framework for the organization of project (process-related as well as product-related) in construction projects. Its purpose is to facilitate control, exchange, retrieval and use of relevant information about the project and the construction entity. It is intended for all agents in the project organization in management of the construction process as a whole and in coordination of its sub-processes and activities.	None	https://www.iso.org/standard/40835.html
ISO 13567-1:2017	Technical product documentation. Organization and naming of layers for CAD. Overview and principles	The International Organization for Standardization (ISO)	2017	Information producers	Structuring and Naming Digital Architectural Data	Specification for structuring layers within computer aided design (CAD) files	It establishes the general principles of layer structuring within CAD files.	None	https://www.iso.org/obp/ui/#iso:std:iso:13567:-1:ed-2:v1:en
ISO 13567-2:2017	Technical product documentation. Organization and naming of layers for CAD. Concepts, format and codes used in construction documentation	The International Organization for Standardization (ISO)	2017	Information producers	Structuring and Naming Digital Architectural Data	Specification for structuring layers within computer aided design (CAD) files	It covers the organization and allocation of layers for CAD on construction projects for the purposes of communication and management.	ISO 13567-1	https://www.iso.org/obp/ui/#iso:std:iso:13567:-2:ed-2:v1:en
BS 8541-1:2012	Library objects for architecture, engineering and construction. Identification and classification. Code of practice	British Standards Institute (BSI)	2012	Information producers	Structuring and Naming Digital Architectural Data	Code of practice for the naming of objects within an object library		BS 1192:2007, BS ISO 10303-21, ISO 10303-28, ISO/PAS 16739, BS EN ISO 9431	https://shop.bsigroup.com/ProductDetail?pid=000000000030240255

BS 8541-4:2012	Library objects for architecture, engineering and construction. Attributes for specification and assessment. Code of practice	British Standards Institute (BSI)	2012	Information producers	Structuring and Naming Digital Architectural Data	Code of practice for specifying the attributes of library objects	Gives recommendations for the application of construction objects integrated into BIM working, defining the level of information for specific uses including specifying desired outcomes and selection of products. Covers common and specific attributes and can be used for the assessment of expected impacts.	BS 8541-1:2012, BS ISO 80000-1	https://shop.bsigroup.com/ProductDetail?pid=000000000030240260
BS 8541-5:2015	Library objects for architecture, engineering and construction. Assemblies. Code of practice	British Standards Institute (BSI)	2015	Information producers	Data Management and Sharing	Code of practice for the transmitting of assemblies of library objects	Provides guidance and recommendations for sharing sub-models of combinations of components and spaces, covering naming, classification and nesting.	BS 8541-1, BS 8541-2, BS 8541-3, BS 8541-4	https://shop.bsigroup.com/ProductDetail/?pid=000000000030294737
ISO 16739-1:2018	Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries	The International Organization for Standardization (ISO)	2018	Software vendors	Data Management and Sharing	Standards for construction information using the IFC schema	The Industry Foundation Classes, IFC, are an open international standard for Building Information Model (BIM) data that are exchanged and shared among software applications used by the various participants in the construction or facility management industry sector. The standard includes definitions that cover data required for buildings over their life cycle. This release, and upcoming releases, extend the scope to include data definitions for infrastructure assets over their life cycle as well.	ISO 10303-11, ISO 10303-21, ISO 10303-28	https://www.iso.org/standard/70303.html
ISO 12006-2:2015	Building construction. Organization of information about construction works. Framework for classification	The International Organization for Standardization (ISO)	2015	Software vendors	Data Management and Sharing	Framework for the development of built environment classification systems.	Framework for the development of built environment systems. It identifies a set of recommended classification table titles for a range of information object classes according to particular views, e.g. by form or function, supported by definitions. It shows how the object classes classified in each table are related, as a series of systems and subsystems, e.g. in a building information model.	ISO 22274	https://www.iso.org/standard/61753.html
BS ISO 16354:2013	Guidelines for knowledge libraries and object libraries	The International Organization for Standardization (ISO)	2013	Software vendors	Data Management and Sharing	Guidelines for the structuring of knowledge libraries and object libraries	Categorize knowledge libraries and object libraries and provide recommendations for the creation of such libraries. Libraries that are compliant with the guidelines of ISO 16354:2013 may be more easily linked to, or integrated with other libraries	None	https://www.iso.org/standard/56434.html
NBSIR 80-1978	Digital Representation for Communication of Product Definition Data (IGES)	National Bureau of Standards (NBS)	1978	Software vendors	Data Management and Sharing	Specification for the exchange of data between Computer Aided Design/Computer Aided Manufacturing	After the initial release of STEP (ISO 10303) in 1994, interest in further development of IGES declined, and Version 5.3 (1996) was the last published standard.	NBSIR 80-1978 ASME Y14.26 M	https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nbsir80-

						(CAD/CAM) systems.			1978.pdf
ISO 16757-1:2015	Data structures for electronic product catalogues for building services. Part One: Concepts, architecture and model	The International Organization for Standardization (ISO)	2015	Software vendors	Data Management and Sharing	Schema for product catalogues to transmit information relating to building services products	Its purpose it the provision of data structures for electronic product catalogues to transmit building services product data automatically into models of building services software applications. This includes a meta model for the specification of product classes and their properties and a meta model for the product data which is exchanged in product catalogues. Product data has to follow the specifications for their product groups.	None	https://www.iso.org/standard/57613.html
ISO 12006-3:2007	Building construction. Organization of information about construction works. Framework for object-oriented information	The International Organization for Standardization (ISO)	2007	Software vendors	Data Management and Sharing	Specification for data dictionaries used to store construction information	It specifies a language-independent model which can be used for the development of dictionaries used to store or provide information about construction works. It enables classification systems, information models, object models and process models to be referenced from within a common framework.	ISO 10303-11, ISO/IEC 10646	https://www.iso.org/standard/38706.html
BS 1192-4:2014	Collaborative production of information. Fulfilling employers information exchange requirements using COBie. Code of practice	British Standards Institute (BSI)	2014	Software vendors	Data Management and Sharing	A code of practice for exchanging handover information back to an employer	It provides methodology for the transfer between parties of structured information relating to Facilities, including buildings and infrastructure. It defines expectations for the design and construction project phases prior to handover and acquisition and the subsequent in-use phase. This code of practice assists the demand side, including employers with portfolio managers, asset managers and facility managers, to specify their expectations while helping information providers, including the lead designers and contractors,to prepare concise, unambiguous and accessible information.	BS 1192:2007, PAS 1192-2, PAS 1192-3, BS ISO 16739	https://shop.bsigroup.com/ProductDetail?pid=0000000003029467 2
ISO 10303	Industrial automation systems and integration -- Product data representation and exchange (STEP)	The International Organization for Standardization (ISO)	varies	Software vendors	Data Management and Sharing	(STEP) The Standard for the Exchange of Product Model Data, is a comprehensive ISO standard (ISO 10303) that describes how to represent and exchange digital product information.		None	Has numerous parts, which can be found here: https://www.iso.org/committee/54158/x/catalogue/p/1/u/0/w/0/d/0
ISO/TR 16310:2014	Symbol libraries for construction	The International	2014	Software vendors	Drawings and Constructio	Requirements for standardized	Specifies the requirements and needs for supplying and managing standardized	None	https://www.iso.org/st

	and facilities management	Organization for Standardization (ISO)			n Documentation	symbolic descriptions of objects	symbolic descriptions of objects that need to be specified in the construction process. Within this context, the term "symbol" is interpreted to cover pure symbolic presentation as well as simplified representation of geometrical shapes of objects.		andard/56142.html
ISO 16757-2:2016	Data structures for electronic product catalogues for building services. Part Two: Geometry	The International Organization for Standardization (ISO)	2016	Software vendors	Data Management and Sharing	Schema for product catalogues to transmit geometry relating to building services products	It describes the modelling of building services product geometry. The description is optimized for the interchange of product catalogue data and includes shapes for representing the product itself, symbolic shapes for the visualization of the product's function in schematic diagrams, spaces for functional requirements, surfaces for visualization, and ports to represent connectivity between different objects.	None	https://www.iso.org/standard/62080.html
ISO 15489-1:2016	Information and documentation – Records Management: Part 1, Concepts and principles	The International Organization for Standardization (ISO)	2016	Custodial Institution	Data Management and Sharing	Archival records management best practice	Concepts and principles from which approaches to the creation, capture and management of records are developed		https://www.iso.org/standard/62542.html
ISO 19005-4:2020	Document management — Electronic document file format for long-term preservation — Part 4: Use of ISO 32000-2 (PDF/A-4)	The International Organization for Standardization (ISO)	2020	Information producers	Representation conventions	Archival profile for PDF	use of the Portable Document Format (PDF) 2.0, as formalized in ISO 32000-2:—, for preserving the static visual representation of page based electronic documents over time in addition to allowing any type of other content to be included as an embedded file or attachment		https://www.iso.org/standard/71832.html
ISO/IEC 26300-1:2015	Open Document Format for Office Applications	The International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC)	2015	Information producers	Representation conventions	Archival profile for office documents	XML schema for office documents such as includes text documents, spreadsheets, charts and graphical documents like drawings or presentations, but not restricted to these kinds of documents		https://www.iso.org/standard/66363.html
ISO 20652:2006	Space Data and Information Transfer Systems - Producer-Archive Interface - Methodology Abstract Standard	The International Organization for Standardization (ISO)	2006	Information producers / Custodial Institution	Data Management and Sharing	Framework for transfer of digital materials from creator into a digital preservation environment	Methodology for the structure of actions that are required from the initial time of contact between the producer and the archive until the objects of information are received and validated by the archive.		https://www.iso.org/standard/39577.html
ISO 21127:2014	Information and documentation — A	The International Organizat	2014	Information producers	Representation conventions	Guidelines for the exchange of information between cultural	Exchange and integration of heterogeneous scientific documentation relating to museum collections.		https://www.iso.org/standard/

	reference ontology for the interchange of cultural heritage information	ion for Standardization (ISO)				heritage institutions	Formalization of the CIDOC/CRM ontology.		57832.html
ISO 15836-1:2017	Information and documentation — The Dublin Core metadata element set — Part 1: Core elements	The International Organization for Standardization (ISO)	2017	Information producers	Representation conventions	Metadata elements for cross-domain resource description	15 core metadata elements for cross-domain resource description. These terms are part of a larger set of metadata vocabularies maintained by the Dublin Core Metadata Initiative		https://www.iso.org/standard/71339.html
ISO 15836-2:2019	Information and documentation — The Dublin Core metadata element set — Part 2: DCMI Properties and classes	The International Organization for Standardization (ISO)	2019	Information producers	Representation conventions	Metadata elements for cross-domain resource description	vocabulary for cross-domain resource description, known as the Dublin Core metadata terms (hereafter DCMI Metadata Terms). It includes all of the properties and classes in the main namespace of DCMI Metadata Terms		https://www.iso.org/standard/71341.html
ISO 19115-1:2014	Geographic information — Metadata — Part 1: Fundamentals	The International Organization for Standardization (ISO)	2014	Information producers	Representation conventions	Schema for describing geographic information and services	Information about the identification, the extent, the quality, the spatial and temporal aspects, the content, the spatial reference, the portrayal, distribution, and other properties of digital geographic data and services		https://www.iso.org/standard/53798.html
ISO 19118:2011	Geographic information — Encoding	The International Organization for Standardization (ISO)	2011	Information producers	Representation conventions	Encoding rules for use for the interchange of geographic data	requirements for creating encoding rules based on UML schemas, requirements for creating encoding services, and requirements for XML-based encoding rules for neutral interchange of data		https://www.iso.org/standard/44212.html
ISO 19136:2020	Geographic information — Geography Markup Language (GML) — Part 1: Fundamentals	The International Organization for Standardization (ISO)	2020	Information producers	Representation conventions	XML encoding in accordance with ISO 19118 for the transport and storage of geographic information	Transport and storage of geographic information modelled in accordance with the conceptual modelling framework used in the ISO 19100 series of International Standards and including both the spatial and non-spatial properties of geographic features		https://www.iso.org/standard/75676.html
ANSI/NISO Z39.87-2006	Data Dictionary - Technical Metadata for Digital Still Images	American National Standards Institute (ANSI)/National Information Standards Organization (NISO)	2006	Information producers	Representation conventions	Metadata to enable users to develop, exchange, and interpret raster digital image files	Facilitate interoperability between systems, services, and software as well as to support the longterm management of and continuing access to digital image collections		http://www.niso.org/publications/ansiniso-z3987-2006-r2017-data-dictionary-technical-metadata-digital-still-images

ISO 15076-1:2010	Image technology colour management — Architecture, profile format and data structure — Part 1: Based on ICC.1:2010	The International Organization for Standardization (ISO)	2010	Information producers	Representation conventions	Colour profile format	Exchange of information which specifies the intended colour image processing of digital data		https://www.iso.org/standard/54754.html
ISO/TR 22299:2018	Document management — Digital file format recommendations for long-term storage	The International Organization for Standardization (ISO)	2018	Information producers / Custodial Institution	Data Management and Sharing	file formats recomendation for or long-term storage	This document gives guidelines for selecting the most appropriate file format(s) for the storage, usability, and exchange of data with a long-term management objective.	ISO 12651-1	https://www.iso.org/standard/73117.html
BS 10010:2017	Information classification, marking and handling. Specification	British Standards Institute (BSI)	2017	Information producers	Data Management and Sharing	security categories and requirements	It sets out guidance for organizations to consistently think about the sensitivity of the information assets that they create, use and share; as well as how to handle them appropriately.	None	https://shop.bsigroup.com/ProductDetail?pid=000000000030330941
IEC 82045-1:2001	Document management — Part 1: Principles and methods	The International Organization for Standardization (ISO)	2001	Information producers / Custodial Institution	Data Management and Sharing	principles for defining metadata for the management of documents	Principles and methods to define metadata for the management of documents associated with objects throughout their life cycle; This cycle generally covers a range from the conceptual idea of a document to its deletion. The established principles and methods are basic for all document management systems.	IEC 61346-1:1996; IEC 61355:1997; IEC 2023:2000; ISO/IEC 2382-1:1993; ISO/IEC 8613-1:1994; ISO 9000:2000; ISO 10007:1995; ISO 15226:1999; ISO 16016:2000	https://www.iso.org/standard/34159.html
IEC 82045-2:2004	Document management – Part 2: Reference collection of metadata and reference models	The International Electrotechnical Commission (IEC)	2004	Information producers / Custodial Institution	Data Management and Sharing	standardized metadata elements for document management.	A set of standardized metadata elements for document management in accordance with IEC 82045-1.	IEC 61346 (all parts); IEC 82045-1:2001; ISO 31-0:1992; ISO 639-1:2002; ISO/IEC 2382-1:1993; ISO 3166-1:1997; ISO 5455:1979; ISO 5457:1999; ISO 8601:2000;ISO 10303- 1:1994,	https://www.iso.org/standard/34513.html
ISO 30300:2020	Information and documentation — Records management — Core concepts and vocabulary	The International Organization for Standardization (ISO)	2020	Information producers / Custodial Institution	Data Management and Sharing	vocabulary and definitions relevant to the core concepts of the records management	This document contains terms and definitions that are relevant to the core concepts of the records management domain. It does not limit the definition of new terms in ISO/TC 46/SC 11 standards.	None	https://www.iso.org/obp/ui/#iso:std:iso:30300:ed-2:v1:en

The compilation of this list would not have been possible without the help of the Building for Tomorrow Standards Working Group that included experts and stakeholders from different fields. Special thanks to John Bacus; Leif Granholm, Nancy Hadley, Leslie Johnston, Marek Suchocki, and Dan Rossiter.

Appendix 8: Principles Resources

This is available online at the Building for Tomorrow website.

Good Practices Resources

Reference	Title	ource	ar	ementer / y Audience	tegories	nplified Scope	Scope and Notes	Normative Reference (Documents that are indispensable for the application of this document)	Reference URL
National Digital information Infrastructure and Preservation Program NDIIPP:	Recommended Formats Statement	Library of Congress		Custodial Institution/ Information producers	File Formats		Recommended Formats Statement identifies hierarchies of the physical and technical characteristics of creative formats, both analog and digital, which will best meet the needs of all concerned, maximizing the chances for survival and continued accessibility of creative content well into the future.		https://www.loc.gov/preservation/resources/rfs/
Community-owned digital preservation tool registry	COPTR	Aligning National Approaches to Digital Preservation initiative			Access Practices	Tools for long-term digital preservation	Finding and evaluation tool to help practitioners discover the tools they need to perform particular preservation tasks.		https://coptr.digipres.org/Main_Page
Cyndi Shein	From Accession to Access: A Born-Digital Materials Case Study	Getty Research Institute	2014	Art Institution	Accession + Access Practices		Between 2011 and 2013 the Getty Institutional Records and Archives made its first foray into the comprehensive ingest, arrangement, description, and delivery of unique born-digital material when it received oral history interviews generated by some of the Pacific Standard Time: Art in L.A. project partners. This case study touches upon the challenges and affordances inherent to this hybrid collection of audiovisual recordings, digital mixed-media files, and analog transcripts. It describes the Archives' efforts to develop a basic processing workflow that applies the resource-management strategy commonly known as "MPLP" in a digital environment, while striving to safeguard the integrity and authenticity of the files, adhere to professional standards, and uphold		https://digitalcommons.usu.edu/westernarchives/vol5/iss1/1/

							fundamental archival principles. The study describes the resulting workflow and highlights a few of the inexpensive technologies that were successfully employed to automate or expedite steps in the processing of content that was transferred via easily-accessible media and consisted of current file formats		
Paradigm	PARADIGM Workbook on Digital Private Papers	Bodleian Library, John Rylands University Library, University of Oxford (no longer in print)	Data Collected 2005-2007, Published 2019		Appraisal Practices	best practice workbook for accessioning and ingesting digital private papers into digital repositories, and processing these in line with archival and digital preservation requirements.	The Personal Archives Accessible in Digital Media (PARADIGM) project saw the major research libraries of the Universities of Oxford and Manchester come together to explore the issues involved in preserving digital private papers through gaining practical experience in accessioning and ingesting digital private papers into digital repositories, and processing these in line with archival and digital preservation requirements. The project outcomes included best-practice guidelines in the form of a workbook on issues relating to the archiving of personal papers in digital form		https://ora.ox.ac.uk/objects/uuid:116a4658-deff-4b06-81c5-c9c2071bc6d0
Michael J. Shallcross, Christopher J. Prom	Appraisal + Acquisitions Strategies	SAA			Appraisal Practices		pp 160-161;		https://scholar.google.com/scholar?q=Appraisal%20and%20Acquisition%20Strategies%20.%20edited%20by%20Michael%20Shallcross%20and%20Christopher%20J.%20Prom.%20Chicago:%20Society%20of%20American%20Archivists,%202016.%20Ap

									praisal%20and%20Acquisition%20Strategies%20Shall cross%20Michael%20Prom%20Christopher%20J.%20Chicago%20Society%20of%20American%20Archivists%202016
Archivists at University of British Columbia	interPARES	Social Sciences and Humanities Research Council of Canada's Community-University Research Alliance (SSHRC-CURA)	1999-2018		Best Practices	interPARES resource for best practices and papers relating to digital preservation in these categories: security, infrastructure, legal, social/societal issues, education, policy, access, control, and terminology	The International Research on Permanent Authentic Records in Electronic Systems (InterPARES) aims at developing the knowledge essential to the long-term preservation of authentic records created and/or maintained in digital form and providing the basis for standards, policies, strategies and plans of action capable of ensuring the longevity of such material and the ability of its users to trust its authenticity. The findings and products of the first three phases of the project can be found on this website.		http://www.interpares.org/
Archivists at University of British Columbia	interPARES TRUST	Social Sciences and Humanities Research Council of Canada's Community-University Research Alliance (SSHRC-CURA)	2019-2020		Best Practices	4th Installment of InterPARES, papers published include categories of security, infrastructure, legal, social/societal issues, education, policy, access, control,and terminology	multi-national, interdisciplinary research project exploring issues of trust and trustworthiness of records and data in online environments. Its goal is to generate the theoretical and methodological frameworks to develop local, national and international policies, procedures, regulations, standards and legislation, in order to ensure public trust grounded on evidence of good governance, a strong digital economy, and a persistent digital memory.		https://interpares trust.org/trust/about_research/studies

	Digital Preservation Handbook	Digital Preservation Coalition (DPC)	2015	Custodial Institution	Best Practices	Knowledge base for digital preservation	Digital information is increasingly important to our culture, knowledge base and economy. The Handbook, first compiled by Neil Beagrie and Maggie Jones in 2001, is maintained and updated by the DPC. This full revision (the 2nd Edition) has expanded and updated content to cover over 30 major sections (see Contents). The 2nd edition was compiled with input from 45 practitioners and experts in digital preservation under the direction of Neil Beagrie as managing editor and William Kilbride as chair of the Management and Advisory Boards. The Handbook provides an internationally authoritative and practical guide to the subject of managing digital resourcesover time and the issues in sustaining access to them. It will be of interest to all those involved in the creation and management of digital materials.		https://www.dpconline.org/handbook
Gore, Emily & Keller, Paul, eds.	Recommendations for Standardized International Rights Statements	Digital Public Libary of American, Europeana	2018	Custodial Institutions/ Information producers	Best practices		Rights Statements to communicate to users the copyright and related restrictions on use of Items in their collections		https://rightsstatements.org/files/180531recommendations_for_standardized_international_rights_statements_v1.2.2.pdf
Michael Boock, Brian E Davis	Next Steps for Building a Flexible and Robust Digital Preservation Infrastructure at Oregon State University Libraries & Press	State University of Oregon	2017		Digital Preservation Infrastructure		2 year plan		https://ir.library.oregonstate.edu/concern/technical_reports/73666904w
Digital Library Federal (DLF)	National Digital Stewardship Alliance studies (NDSA)	Library of Congress, National Digital Information Infrastructureand Preservation Program	2010		Digital Storage	Library of Congress provided secretariat and membership management services to NDSA members			

Yan Han	Beyond TIFF and JPEG2000: PDF/A as an OAIS submission information package container	University of Arizona Libraries	2015		File Formats				
Caroline R. Arms, Carl Fleischhauer, Kate Murray, Marcus Nappier	Library of Congress Sustainability of Digital Formats	Library of Congress	started in 2004	Custodial Institution/ Information producers	Format Preservation	Information about digital content formats through detailed format description documents or fdds	This site is devoted to the analysis of the technical aspects of digital formats.This analysis will inevitably have implications for policy matters, most significantly collection policies. It is concerned with the formats associated with media-independent ("intangible") digital content, i.e., content that is typically managed as files and which is generally not dependent upon a particular physical medium. It is not concerned with the formats associated with media-dependent ("tangible") digital content, i.e., formats that are dependent upon and inextricably linked to physical media, e.g., DVDs, audio CDs, and videotape formats like DigiBeta.	None	https://www.loc.gov/preservation/digital/formats/
Public Record Office, Royal Commission on Historical Manuscripts, Her Majesty;'s Stationery Office, Office of Public Sector Information	PRONOM	National Archives, UK	2011	Custodial Institution	Format Preservation	Information about software products, and the file formats which each product can read and write	Resource for anyone requiring impartial and definitive information about the file formats, software products and other technical components required to support long-term access to electronic records and other digital objects of cultural, historical or business value.	None	https://www.nationalarchives.gov.uk/PRONOM/Default.aspx
Archivist of US: David S. Ferriero	U.S. National Archives and Records Administration Digital Preservation Framework	NARA	est 1934	Custodial Institution/ Information producers	Format Preservation	Risk and Prioritization Matrix and File Format Preservation Action Plans.		None	https://github.com/usnationalarchives/digital-preservation
Katherine Thornton, Kenneth Seals-Nutt, Euan Cochrane, Carl Wilson	Wikidata for Digital Preservation		2006	Custodial Institution	Format Preservation	Technical registry of metadata related to computer software and computing environments	The vision of Wikidata for Digital Preservation is to use Wikidata as a technical registry of metadata related to computer software and computing environments. Collaboratively creating this metadata, and making it available as linked open data, will reduce the amount of redundant work digital preservation professionals do in		https://www.wikidata.org/wiki/Q31742273

							order to describe resources.		
	Library of Congress Sustainability of Digital Formats	Library of Congress		Custodial Institution	Format Preservation	Information about digital content formats through detailed format description documents or fdds	This site is devoted to the analysis of the technical aspects of digital formats. This analysis will inevitably have implications for policy matters, most significantly collection policies. It is concerned with the formats associated with media-independent ("intangible") digital content, i.e., content that is typically managed as files and which is generally not dependent upon a particular physical medium. It is not concerned with the formats associated with media-dependent ("tangible") digital content, i.e., formats that are dependent upon and inextricably linked to physical media, e.g., DVDs, audio CDs, and videotape formats like DigiBeta.		https://www.loc.gov/preservation/digital/formats/
collaboration with British library, library of congress, harvard library, NARA, DPC	British Library File Format Assessments	British Library		Custodial Institution/ Information producers	Ingest Practices	a file format assessment of best practices	File format assessments capturing knowledge about the gaps in current best practice, understanding and capability in working with specific file formats. The focus of each assessment is on capturing evidence-based preservation risks and the implications of institutional obsolescence which lead to problems maintaining the content over time.		https://wiki.dpconline.org/index.php?title=File_Formats_Assessments

Tools

Tool	Source	Category	Open Source (y/n)	Description	Link
Box		Cloud storage provider	n	cloud-hosted service provider for entire file lifecycles - creation, storage, co-editing, signature, etc.	https://www.box.com/overview
Amazon Glacier		Commercial cloud Storage offerings	n	fee per service, upload, fixity check, copies made, long term storage, not a full digital repository, does not deal with metadata - different from s3 is cost and service level - appropriate for dark archiving, no public access and no fixed time for retrival, may take2 days to receive requested item, costs less, not appropriate if looking for fast and many requests	https://aws.amazon.com/glacier/
Amazon S3		Commercial cloud Storage offerings	n	fee per service, upload, fixity check, copies made, long term storage, not a full digital repository, does not deal with metadata - costs more than glacier but immediate access to items uploaded, appropriate fortransactional material or if something needs to be publicly available	https://aws.amazon.com/s3/
Internet archiva		Online Service	y	Free online service that attempts to save and make accessible content	https://archive.org/
Archive-It		Online service + storage	n	hosted preservation storage service, and access to web archives (capture, storage, + delivery, there is apilot project for preservation services available for any kind of content	https://archive-it.org/
DuraCloud	Lyrasis (formerly Duraspace)	Online storage platform	n	hosted preservation storage service, can be a component of a larger preservation program (lyrasis)	https://duraspace.org/duracloud/
MetaArchive Cooperative		Organization - Membership Service	n	provides commonly supported preservation servicesfor members, uses LOCKSS and made available preservation materials for members	https://metaarchive.org/
EaaSI (emulation as aservice infrastructure)	Yale Library	Project	y	project producing open source software product, made specifically for digital software preservation - for preservation and access in emulated environment	https://educopia.org/emulation-as-a-service-eaasi/#:~:text=2017%2D2020-,Scaling%20Emulation%20as%20a%20Service%20Infrastructure%20(EaaSI)%20(subcontract),inherent%20obsolescence%20of%20developing%20technologies.

BitCurator		Software	Y	Provides digital forensics capability, associatedcommunity: BitCurator Consortion (formal membership)	https://bitcurator.net/
Archivematica		Software	Y	digital preservation repository platform, equivalent to RDRS, archival community focused but applicable toother types of repositories	https://www.archivematica.org/en/
Preservica		Software	N	digital preservation repository, available in two modes: locally stored or hosted version	https://preservica.com/
Forensic Toolkit		Software	Y	Similar to BitCurator	https://accessdata.com/products-services/forensic-toolkit-ftk
Bagger	Library of Congress & California Digital Library	Software	Y	open source application that packages data files according to the BagIt specification. Add BagIt infohere	https://www.loc.gov/preservation/digital/
CONTENTdm	OCLC	Software		content management system; main audience: content manager and curators, build and showcase your digital collections on your personalized website, making them more discoverable to people around theworld. CONTENTdm also secures and monitors your digital originals in a cloud-based preservation archive so they remain safe for the future.	https://www.oclc.org/en/contentdm.html
DSpace	MIT	Software	Y	digital repository, can be used for preservation, but often used as institutional repository, optimized for self-service use, A specialized content management system that allows different communities to use the web to capture, distribute, and preserve digital worksand to provide access to those works through metadata.	http://www.dspace.org/
Islandora		Software	Y	Respository/content management system from Fedora + Drupal (content management system for curatorial and user access)	https://islandora.ca/
ArchivesSpace		Software	Y	comprehensive software workbench for archivalmaterial processing (digital and non-digital); acquisition, appraisal, etc. primary audience: archivists, very common	https://archivesspace.org/
DROID	National Archives UK	Software	Y	File format identification	https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/
FTK Imager		Software		digital forensics, disk image - method to create a file that represents entire digital media and put into astandardize structure for examination	https://accessdata.com/products- services/forensic-toolkit-ftk/ftkimager
Rosetta		Software	N	Commercial vended software product from EX libris, digital repository product, commercial alternative tofedora and	https://exlibrisgroup.com/products/rosetta-digital-asset-management-and-preservation/

				archivematica	
Microsoft File Checksum Integrity Verifier		Software	y	Confirms file fixity	https://support.microsoft.com/en-us/topic/d92a713f-d793-7bd8-b0a4-4db811e29559
Luna		Software	n	Commercial software project for managing digital image collections	http://www.lunaimaging.com/
Aid4Mail		Software	n	used within email archiving programs, tool that translates between different email formats (e.g., PST,EML, etc)	https://www.aid4mail.com/
FITS (File Information Tool Set)	Harvard Library	Software	y	identifying and validating file formats, extracting metadata embedded within files, and outputting the metadata in various formats. It was created to do some of the file processing tasks needed to supportdigital preservation repositories and applications.	https://projects.iq.harvard.edu/fits
BagIt	Library of Congress & California Digital Library	Specification of File Packaging	y	a hierarchical file packaging format for the creation of standardised digital containers called 'bags,' which are used for storing and transferring digital content.	https://www.loc.gov/preservation/digital/
Data Accessioner	Duke University Rare Book, Manuscripts, and Special Collections Library	Software	y	primarily used for migrating content between media,creating and validating checksums, gathering metadata, compiling an XML metadata file for future reference	http://dataaccessioner.org/