# Archiving digital architectural records: towards a national framework

## PROJECT REPORT

## June 2018

**Prepared by Chris Burns**

**School of Art, Architecture and Design**

**University of South Australia**

University of South Australia

ARCHITECTURE MUSEUM

NATSPEC// Construction Information

## Contents Page

## Acknowledgements

The project team gratefully acknowledges the assistance of the following:

**NATSPEC**

Richard Choy, CEO

Jocelyn Holley, Development Manager

**University of South Australia**

Darren Fong, Lecturer, University of South Australia

Steve Kelly, Lecturer, University of South Australia

Damian Madigan, Lecturer, University of South Australia

Jeffrey Dunn, Information Technologist, University of South Australia

**National Library of Australia**

Beth Lonergan, Curator of Digital Archives (a/g), Pictures and Manuscripts Branch, National Library of Australia

Libor Coufal, Assisstant Director, Digital Preservation; Director, Preservation Services (acting), National Library of Australia

**Individuals**

Harriet Edquist, Director, RMIT Design Archives

Rickie-Lee Robbie, Collections Database Coordinator, RMIT Design Archives

on behalf of

Katrina Chisholm, Registered Architect, Nedlands, Western Australia

Colum O'Riordan, CEO, Irish Architectural Archive

This project was funded by NATSPEC.

# 1. Introduction

## 1.1 Project Aims and Objectives

### 1.1.1 The Project

This project is named *Archiving digital architectural records: towards a national framework* (ADAR:NF). It builds upon the foundations laid by a 2015-16 project in which members of the project team had leading roles. The previous project, *Securing and enabling access to knowledge for the future: archiving digital architectural records* (ADAR) included a two-day public Symposium, *Born digital: a symposium exploring digital architectural and built environment records,* (18-19 April 2016). One of the recommendations of the symposium was that guidelines for archiving digital architectural records be developed for use by archivists and architects.

Mr Richard Choy, CEO of NATSPEC, attended and spoke at the 2016 symposium. NATSPEC is a national not-for-profit organisation, owned by Government and industry, whose objective is to improve the construction quality and productivity of the built environment through leadership of information. NATSPEC generously offered to provide funding to enable further research into current digital archiving practices (via an online questionnaire and literature review) and the preparation of a draft national framework for archiving digital architectural records. The funding allowed the research team to employ a research assistant, Mr Chris Burns. The project concluded in June 2018.

### 1.1.2 Overview

The emergence of digital technology has had a significant impact on the way in which buildings are designed and constructed. From being regarded initially as a tool to aid design, the computer is now commonly considered to be integral to the design process. The digital environment in which an architectural project is developed involves computer hardware and software in the creation of digital files.

The records of the process of designing a building cover a broad spectrum. They include models that explore its potential shape and form, sketches, plans, elevations, sections, renderings and other documents like photographs, emails, letters, faxes, specifications and contracts. Increasingly, these records are produced in digital environments and only exist as digital files.

Currently there is no Australian national framework to guide collecting, archiving and preserving these "born digital" architectural records (however, investigations are underway overseas). As a consequence, their future accessibility to practitioners and researchers is in jeopardy.

The key challenges associated with archiving born digital architectural records include: (1) the rationale for collecting records produced in a digital environment; (2) which records to archive; (3) how to archive them; and (4) how to achieve digital continuity in rapidly evolving and changing electronic environments.

#### 1.1.3 Scope

The project:

1. Undertook a scoping literature review and online questionnaire focussed on the four challenges identified in 1.1.2.
2. Drafted a framework for archiving born digital architectural records. The framework was tested with a limited number of selected stakeholders. The framework is intended to be usable for digital architectural records and transferrable to other disciplinary categories, e.g. in building and construction and engineering.

#### 1.1.4 Outputs

The project led to several outputs:

1. A literature review, summarising existing literature and identifying existing guidelines for archiving born digital architectural records.
2. A draft national framework for a) archival professionals and b) architectural practitioners based on the literature review and the survey results.
3. Project report.
4. Project findings will be disseminated through peer-reviewed and professional journal articles.

## 1.2 Project Team

### Assoc Prof Christine Garnaut

Christine Garnaut is a planning and architectural historian and the inaugural Director of the Architecture Museum (formerly the Architecture Archive) at UniSA. The Museum collects records of South Australian-based architects and related professionals. It is Australia's only Architecture Museum and a member of the International Confederation of Architectural Museums (ICAM). Christine Garnaut is an ICAM Board member and Convenor of the regional network ICAM Australasia. E: Christine.Garnaut@unisa.edu.au

### Dr Julie Collins

Julie Collins holds a B. Arch. and a PhD in architecture. She is Collections Manager and researcher at the Architecture Museum, School of Art Architecture and Design. She has extensive knowledge including of best practice in architectural archival records management. She provides advice to the architecture profession about how to manage their hardcopy records. E: Julie.Collins@unisa.edu.au

### Mr Chris Burns

Chris Burns is the Research Assistant for the project *Archiving digital architectural records: towards a national framework*. Previously, Chris was Research assistant for the projects *Securing and enabling access to knowledge for the future: archiving digital architectural records*, upon which this current project builds. He holds a B. Industrial Design and a M. Education (Design and Technology) from the University of South Australia. E: Chris.Burns@unisa.edu.au

## 1.3 Partners

**NATSPEC** is a national not-for-profit organisation, owned by Government and industry, whose objective is to improve the construction quality and productivity of the built environment through leadership of information. https://www.natspec.com.au/

**icam Australasia** is a regional network of the International Confederation of Architectural Museums (icam).

The **Architecture Museum,** School of Art, Architecture and Design, University of South Australia, is a nationally unique repository of architects' and allied professionals' records and a dynamic hub of research into South Australia's architectural and built environment history.

## 1.4 Methods

### 1.4.1 Literature Review

A literature review expanded on research conducted during our earlier ADAR project. It served to summarise literature to date and identified existing guidelines for archiving born digital architectural records (see section 2).

### 1.4.3 Online Questionnaire

An online questionnaire provided the researchers with insights into the status of professional archivists' (libraries, archives, museums) and architectural practitioners' endeavours to archive born digital records.

Core recruitment used a database containing the contact details of icam Australasia network members. The Chief Investigator invited selected members to participate via email.

The questionnaire contained two sets of questions; the first tailored for individuals based in architectural practices, the second for individuals based in collecting institutions (libraries, archives, museums). Each participant answered one set of questions only, as the survey branched after the first page, depending on the response to an initial sorting question.

The questionnaire contained a number of closed and open questions, and was anonymous. However, a text box on the final page provided participants with the opportunity to enter their contact details if they were willing to provide additional clarification or further information on any of their responses.

Prior to recruiting participants, an application was submitted to the University of South Australia's Human Research Ethics Committee. The application was approved and assigned protocol 200996.

The online questionnaire was conducted in accordance with the approved application.

## 2. Literature Review

The key challenges associated with archiving born digital architectural records include: (1) the rationale for collecting records produced in a digital environment; (2) which records to archive; (3) how to archive them; and (4) how to achieve digital continuity in rapidly evolving and changing electronic environments. The literature associated with outlining and discussing these challenges is broad, and is related to areas including architectural practice, archival theory and practice, and digital preservation.

This literature review draws upon recent literature to address the following key themes:

- Long-term digital preservation
  - Bitstream preservation and fixity
  - Digital repositories and legacy physical carriers
  - Emerging 'archival' carrying media
- Selected digital preservation standards and guidelines
- Special challenges of preserving 3D CAD models
- Issues surrounding digital architectural records in architectural practice
  - Early born digital architectural records
  - Naming Conventions
- Issues surrounding digital architectural records in collecting institutions
  - Digital collection policies
  - Donor guidelines
  - Metadata

### 2.1 Long-term digital preservation

The affordance of lossless copying confers potential immortality on digital objects. However, with a few potential exceptions (discussed below), all commercially available digital carrying media are subject to both slow degradations over time and the possibility of sudden, catastrophic failure.[1] Therefore, while traditional, paper-based archives may survive benign neglect over decades or even centuries, the records held in digital repositories require ongoing maintenance and active stewardship to preserve their integrity.

Accordingly, the expertise necessary to achieve long-term preservation of digital objects is different from that required in traditional archival practice. Douglas Elford, Lisa Jeong-Reuss, Somaya Langley and Melanie Wilkinson from the National Library of Australia (NLA) discuss commonalities and differences, drawing a contrast between the innovation and experimentation that characterises the emergent field of digital preservation and the tried-and-tested methodologies and techniques employed in traditional conservation.[2]

---

[1] David Rosenthal, Thomas Robertson, Tom Lipkis, Vicky Reich, Seth Morabito, 'Requirements for digital preservation systems: a bottom-up approach' in *D-Lib Magazine* 11, No. 11, www.dlib.org/dlib/november05/11rosenthal.htm accessed 28 May 2018

[2] Douglas Elford, Lisa Jeong-Reuss, Somayal Langley and Melanie Wilkinson, *Getting the whole picture: finding a common language between digital preservation and conservation* (2012) https://www.nla.gov.au/content/getting-the-whole-picture-finding-a-common-language-between-digital-preservation-and accessed 28 June 2018

Michelle Lindlar and Hedda Saemann identify three digital object layers, adapted loosely from the work of Kenneth Thibodeau,[3] which offer a useful framework for contextualising conversations about the preservation of digital objects:

- The bitstream layer is associated with the physical storage of binary code (usually represented as zeroes and ones) that comprises digital objects. At the bitstream level, the integrity of digital objects may be compromised by degradation over time (bit rot), copying errors, catastrophic hardware failure or accidental deletion, while access to digital objects may be impeded by obsolescence of physical carrying media.
- The logical layer refers to the human-interpretable content of digital files. At the logical layer, access to digital objects may be impeded by a dependency on proprietary file formats and/or software. Digital files themselves may also be inadvertently modified or overwritten.
- The semantic layer refers to human understandings of and interactions with the content of digital files. At the semantic layer, the interpretation of digital content may be dependent on context or provenance, while specialised, tacit knowledge may be required to interact with digital files in complex software environments.[4]

#### 2.1.1 Bitstream preservation and fixity

Bitstream preservation is the foundation of digital preservation; if the integrity of the underlying bitstreams of digital files is not preserved, all other preservation actions are in vain.

At the bitstream level, preservation of all digital files, regardless of format or complexity, means replication[5] – maintaining at least two copies of each file in different locations – and monitoring the binary code itself for changes that may occur, for example, randomly in storage or as errors introduced during copying. In practice this is achieved by generating a digital signature for each digital file, known as a checksum; stored in association with each digital object, checksums allow changes in the bit stream to be detected. Once detected, corrupted files must be replaced with undamaged duplicate copies ideally prepared at the time when files were transferred to a digital repository, since the checksum datum does not contain sufficient information to locate and replace the damaged bits. In this sense, bit rot is often irreversible.

Various freely available, open-source software solutions exist both for generating checksums and checking and validating the integrity of digital files against previously generated checksums for complex hierarchies of files and folders.[6]

---

[3] Kenneth Thibodeau, 'Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years' in *The State of Digital Preservation: An International Perspective* (conference proceedings), CLIR Report 07, pp. 4-31, http://www.clir.org/pubs/reports accessed 26 June 2018

[4] Michelle Lindlar and Hedda Saemann, *The DURAARK Project –Long Term Preservation of Architectural 3D Data* (conference proceedings), CIDOC 2014, Annual conference of the International Committee for Documentation/The International Council of Museums, Dresden (2014), http://www.cidoc2014.de/images/sampledata/cidoc/papers/L-1_Lindlar_Saemann_paper.pdf accessed 28 June 2018

[5] Rosenthal et al, 'Requirements for digital preservation systems: a bottom-up approach'

[6] Some examples are provided in Julianna Barrera-Gomez and Ricky Erway, *Walk This Way: detailed steps for transferring born-digital content from media you can read in-house,* Dublin, Ohio: OCLC Research (2013), http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf accessed 4 April 2018

### 2.1.2 Digital repositories and legacy physical carriers

Many guidelines now recommend isolating digital files from disparate carrying media, whether regarded as legacy carriers or otherwise, and ingesting them into a centralised (OAIS-type, see section 2.5) repository where fixity information (checksums) for the entire archive can be generated and checked automatically on a regular basis. For example, Archives New Zealand stresses the importance of separating digital content from physical carriers in their *Digital Preservation Strategy*.[7] Centralising a repository eliminates the need for manually switching between media (changing CD recordable media, for example) which is time consuming and risks damage to physical carriers each time they need to be accessed.

A central repository could be, for example, a server for a well-resourced organisation with good IT support, or outsourced to a third-party provider.[8] For smaller collections of digital content in less well-resourced institutions, a central repository could mean several complete copies of the repository stored on more than one hard drive (preferably three, with one backup offsite) with each periodically checked for integrity.[9]

It is essential that files and their metadata (including, for example, dates created and modified) are not altered during the process of transfer from carrier to repository.[10]

The Canadian Centre for Architecture (CCA) uses write blockers to protect files on carrying media from accidental erasure and creates forensic disk images – 'essentially a 1:1 copy of the data exactly as it resides on the physical media – including any empty space, corrupted sectors, or partially deleted data'[11] – to ensure that the provenance and integrity of digital files is maintained.

Two Online Computer Library Centre, Inc. (OCLC) reports, *You've got to walk before you can run: first steps for managing born-digital content received on physical media,*[12] and *Walk this way: detailed steps for transferring born-digital content from media you can read in-house,*[13] offer, respectively, simple steps and more detailed advice on safely retrieving files stored on physical carrying media. Meanwhile the Digital Preservation Team at the National Library of Australia (NLA) have developed an online, searchable database of physical carriers, called Mediapedia, which may assist in the identification of various physical carrier formats including magnetic tape, optical disks and zip drives.[14]

---

[7] Archives New Zealand, *Digital Preservation Strategy*, http://archives.govt.nz/advice/government-digital-archive-programme/digital-preservation-strategy/digital-preservation-strat accessed 10 May 2018

[8] RLG/OCLC Working Group on Digital Archive Attributes, *Trustworthy Digital Repositories: Attributes and Responsibilities,* Mountain View, California: RLG (May 2002) p. 5, https://www.oclc.org/research/activities/past/rlg/trustedrep/repositories.pdf accessed 28 June 2018

[9] Megan Phillips, Jefferson Bailey, Andrea Goethals and Trevor Owens, *The NDSA Levels of Digital Preservation: An explanation and uses,* Library of Congress, National Digital Stewardship Alliance (2013), www.digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf accessed 16 May 2018

[10] Tim Walsh, *Preservation and Access of Born Digital Architectural Design Records in an OAIS-Type Archive,* (2015) p. 17, https://www.bitarchivist.net/docs/walsh_CADArchiving2015_final.pdf accessed 28 June 2018

[11] Tim Walsh, 'Archaeology of the Digital and Born-Digital Archives at CCA' presentation at *Born digital: a symposium exploring Born digital: a symposium exploring digital architectural and built environment records,* Adelaide, 18 – 19 April 2016: http://www.unisa.edu.au/born-digital accessed 18 June 2018

[12] Ricky Erway, *You've got to walk before you can run: first steps for managing born-digital content received on physical media,* Dublin, Ohio: OCLC Research (2012), http://www.oclc.org/research/publications/library/2012/2012-06.pdf accessed 28 May 2018

[13] Julianna Barrera-Gomez and Ricky Erway, *Walk This Way*

[14] National Library of Australia, "Mediapedia" https://mediapedia.nla.gov.au/home.php accessed 14 June 2018

Tim Walsh, digital archivist at the Canadian Centre for Architecture (CCA), advises that archives always retain a copy of ingested data in its original format, 'as a means of guaranteeing authenticity and to leave options open for future emulation and migration projects.'[15]

Legacy physical carriers may still be retained where appropriate as a last resort backup,[16] but are not capable of carrying the full burden of an organisation's digital preservation efforts.

For further reading, David Rosenthal et al provide extensive, in-depth information on digital preservation systems and their requirements.[17]

**2.1.3 Emerging 'archival' carrying media**

As mentioned, most commercially available digital carrying media have short reliable lifespans and are subject to sudden failure. However, at least two commercial initiatives – one in development (Digital Optical Technology System) (DOTS), another with a product already on the market (M-DISC) – aim to redress this problem by delivering true 'archival' carrying media for the long-term storage of digital information.

Originally developed by Eastman Kodak and now undergoing further development by U.S. based Group 47, DOTS promises stable, write-once storage of digital data in large quantities (a terabyte-capacity DOTS cartridge is pictured in DOTS promotional material available from the Group 47 website).[18] While the DOTS system sounds promising on face value, the technology is still being developed and appears to be some years away from reaching the marketplace.

Meanwhile M-DISC optical media, originally developed by Millenniata, Inc., is currently commercially available and potentially suitable for small-capacity, long-term archiving. Similar in appearance and principle to conventional CD and DVD recordable media, M-DISC incorporates a patented 'rock-like' data layer in lieu of the dye-based or phase-change film data layer which renders standard optical recordable media susceptible to degradation over time.[19] While M-DISC requires special compliant optical drive hardware to write data to disk, once written, the format appears to be largely compatible with (and readable by) most CD/DVD-ROM drives.

M-DISC media underwent extensive environmental testing against conventional optical recordable media by the U.S. Department of Defense in 2009; the final project report noted that 'none of the Millenniata media suffered any degradation at all.'[20]

One limitation of M-DISC DVD media is the small storage capacity of 4.7GB, consistent with conventional single-layer DVD recordable media. M-DISC Blu-ray optical disks have larger capacities ranging from 25 to 125 GB. However, independently-generated longevity data for the Blu-ray variant of this media is not currently available.

---

[15] Tim Walsh, *Preservation and Access of Born Digital Architectural Design Records,* p. 21
[16] Ricky Erway, *You've got to walk before you can run*
[17] David Rosenthal et al, 'Requirements for digital preservation systems: a bottom-up approach'
[18] http://www.group47.com/Group_47-DOTS_Technology_Overview-WEBSITE.pdf accessed 28 May 2018
[19] Fred R. Byers, *Care and Handling of CDs and DVDs: a guide for librarians and archivists,* Washington DC: Council on Library and Information Resources/National Institute of Standards and Technology (2003), http://www.clir.org/pubs/reports/pub/121 accessed 24 June 2018
[20] Ivan Svrcek, *Accelerated Life Cycle Comparison of Millenniata Archival DVD,* China Lake, California: Life Cycle and Environmental Engineering Branch Naval Air Warfare Center Weapons Division (2013).

Storing data on optical media, especially when a large quantity of data is stored over a large number of disks, means fixity checking is more labour-intensive than when files are located in a centralised archive. Nevertheless, M-DISC media may prove suitable for use as an offline, last-resort backup.

## 2.2 Selected digital preservation standards and guidelines

*The Reference Model for an Open Archival Information System (OAIS)* is an international standard (ISO International Standard 14721, updated 1471:2012), developed by the Consultative Committee for Space Data Systems (CCDS) to provide 'a comprehensive and consistent framework for describing and analysing digital preservation issues'.[21] The OAIS model has since become

> almost universally accepted as the lingua franca of digital preservation, shaping and sustaining conversations about digital preservation across disparate domains, and supplying a general mapping of the landscape that stewards of our digital heritage must navigate in order to secure the long-term availability of digital materials.[22]

Brian Lavoie's Digital Preservation Coalition (DPC) Technology Watch Report, *The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition)* offers an accessible and comprehensive introduction to the OAIS reference model, tracing its development, features, benefits and limitations, and its impact on the broader digital archival community.[23]

The OAIS Reference Model introduces the concept of the 'designated community' – the community of primary users of an OAIS-type repository:

> It should not be inferred that the scope of the designated community is determined *ex post* by the nature of the archive's contents; rather, it is the scope of the Designated Community that determines both the contents of the OAIS and the forms in which the contents are preserved, such that they remain available to, and independently understandable by, the Designated Community.[24]

ISO International Standard 16363 describes criteria for the audit and certification of trustworthy digital repositories. The report *Trusted Digital Repositories: Attributes and Responsibilities* (published in 2002 and developed around the then-emerging OAIS Reference Model) defines a 'trusted' repository as one that provides 'reliable, long-term access to managed digital resources to its designated community, now and in the future.'[25]

Drawing on ISO 14721 and 16363, the *Digital Preservation Capability Maturity Model* (DPCMM) describes five levels of digital preservation capability ('maturity stages') across fifteen key areas, benchmarked against ISO 14721 conformance, with metrics against which to measure performance in each area:

---

[21] Brian Lavoie, *The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition),* Digital Preservation Coalition (2014) p. 3, http://dx.doi.org/10.7207/twr14-02 accessed 24 June 2018
[22] Ibid p. 3
[23] Ibid
[24] Ibid p. 10
[25] RLG/OCLC Working Group on Digital Archive Attributes, *Trustworthy Digital Repositories: Attributes and Responsibilities,* Mountain View, California: RLG (May 2002) p. i

> The maturity stages are cumulative: an organisation achieving a higher stage of maturity must implement and sustain all of the requirements for that stage in addition to requirements for all lower stages.[26]

DPCMM was developed to allow archivists 'to conduct a gap analysis of current digital preservation capabilities and to help practitioners and organisations delineate a multi-year roadmap of incremental improvements.'[27]

DPCMM stipulates the adoption of 'open standard technology neutral formats' as preferred preservation formats, and so implementing DPCMM may be problematic when attempting to preserve the significant properties of proprietary 3D CAD formats (see section 2.3). DPCMM provides examples of open standard formats for spreadsheets, text files, 2D images, audio, video, graphics and web pages, but does not mention interactive 3D objects or software which, at the time of writing, are arguably the most problematic file types of all.[28]

The *NDSA Levels of Digital Preservation*, developed by the US-based National Digital Stewardship Alliance, offers a set of plain-language, rubric-based guidelines designed to meet the need for

> straightforward, accessible practices that are more substantial than the conventional advice geared towards individuals, but less daunting and demanding than those required for certification as a trustworthy digital repository.[29]

*Levels of Digital Preservation* contains five preservation categories (Storage and Geographic Location, File Fixity and Data Integrity, Information Security, Metadata and File Formats) and four tiers or levels of increasing sophistication (roughly characterised as 'protect your data', 'know your data', 'monitor your data' and 'repair your data') which, like the DPCMM maturity stages, are cumulative – moving from a 'basic need to ensure bit preservation' towards more advanced requirements.

The *Levels* – 'practical, actionable, and scalable' – are an excellent starting point for 'institutions of all sizes and resource levels' interested in digital preservation.[30]

---

[26] Charles Dollar and Lori Ashley, *Digital Preservation Capability Maturity Model (DPCMM): Background and Performance Metrics Version 2.7* (July 2015) p. 2, https://www.statearchivists.org/resource-center/resource-library/digital-preservation-capability-maturity-model-dpcmm/ accessed 23 May 2018
[27] Ibid, p. 2
[28] Ibid, p. 21
[29] Megan Phillips et al, *The NDSA Levels of Digital Preservation*
[30] Ibid p. 1

## 2.3 Special challenges of preserving 3D CAD models

The first widespread appearance of digital technology in Australian architectural practice occurred in the early 1980s, in parallel with the emergence of the first generation of affordable microcomputers (now usually referred to as desktop computers or PCs). While Computer Aided Design (CAD) software was initially regarded as a replacement for a manual drawing board,

> The move to three dimensions was the point at which CAD models stopped being mere conveniences for drawing blueprints and started taking on importance in their own right. With 3D models, it became possible to design shapes that could not be clearly or adequately expressed by three 2D elevations … In the context of industrial production, 2D surrogates soon became inadequate records and regarded as dangerously open to misinterpretation.[31]

Building Information Models (or BIMs) are an advanced subset of 3D CAD models used primarily in the fields of architecture and engineering, and are described by Michelle Lindlar as 'semantically rich digital objects which contain geometry and layout as well as information on material, cost estimation and scheduling.'[32]

Alex Ball's 2013 DPC Technology Watch Report, *Preserving Computer-Aided Design (CAD)* contains the most comprehensive recent advice and recommendations on the long-term preservation 3D CAD/BIM models.

While preserving the underlying bitstream of native format 3D CAD/BIM models is no different from preserving that of any other file format, preserving access to the intrinsic content of native 3D CAD/BIM models is more difficult than most other file formats (arguably, only software itself is more difficult to preserve). While the logical content of a 2D image, audio file, or video file can be repackaged in an open-source file format without any loss of information or quality (for example, TIFF, WAV, Motion JPEG 2000 respectively),[33] native format 3D CAD/BIM models contain a number of special, functional affordances that usually do not survive conversion into open-source derivatives.

These affordances include:

- Construction history that may be rolled back and forward like an interactive "undo/redo" feature. The history of modelling actions may be rolled back to an earlier point in the model's development, changes implemented, and the model "rolled forward" again, automatically updating later stages of construction with the results of the changes.

---

[31] Alex Ball, *Preserving Computer-Aided Design (CAD)*, (Bath, UK: Digital Preservation Coalition 2013) p. 5, http://dx.doi.org/10.7207/twr13-02 accessed 14 June 2018

[32] Michelle Lindlar, 'Building information modelling – a game changer for interoperability and a chance for digital preservation of architectural data?' presented at iPres2014, 11th International Conference on Digital Preservation, Melbourne (6-10 October 2014) p. 3, https://fedora.phaidra.univie.ac.at/fedora/get/o:378117/bdef:Content/download accessed 24 June 2018

[33] The *Sustainability of Digital Formats: Planning for Library of Congress Collection* website provides sustainability information on various digital file formats: http://www.loc.gov/preservation/digital/formats/fdd/descriptions.shtml accessed 24 June 2018

- Parametric modelling, in which 'aspects of the design are given a variable value instead of a fixed one, in order to make them easier to adjust and reuse in different contexts,' with 'constraints [that] control how designs should be adjusted in the light of changed variables.'[34]
- Features, which are the parametric building blocks of a model within some solid modelling applications (for example, Solidworks and Autodesk Inventor) and in many cases have real-life manufacturing analogues – for example, "draft angle" (taper) added to a moulded part to allow it to be removed from a mould, a "fillet" (curved blended surface) to reduce stress where two surfaces meet, or a solid body hollowed out to create a "shell".

Ball points out that these affordances 'embed far more information into a model than would be evident from just the final shape data … while some information might be inferred using sufficiently advanced feature recognition, having the original information would clearly be preferable if one were using the designs as evidence in a legal case or academic argument.'[35] Thus, much of the information that may be embedded in a CAD/BIM model can only be experienced by opening and interacting with the model in its original, native software environment.

The MIT Libraries FACADE (Future-proofing Architectural Computer-Aided Design) project[36] was an early extensive investigation into the challenges of archiving 3D CAD and BIM models. FACADE made the following broad recommendations for conversion of born-digital architectural records into open formats:

- Manual conversion of 3D CAD models to generate open format derivatives;
- Semi-automated conversion of other 'key design file formats,' for example 2D drawings, into PDF;
- Automated conversion of common digital file formats into open formats, for example 2D documents and images;
- No conversion for other file formats.

Additionally, FACADE recommended that the following versions of 3D CAD models be retained and/or created as appropriate, and retained in perpetuity:

- Original (the originally submitted version of the CAD model)
- Display (an easily viewable format to present to users, normally 3D PDF)
- Standard (full representation in preservable standard format, normally IFC or STEP)
- Desiccated (simple geometry in a preservable standard format, normally IGES)[37]

The authors of the FACADE project final report also noted:

> We believe that it is important to keep the original 3-D model as well, both for authenticity purposes and because most native software is still improving on export capabilities so that it may be possible to create even better standard export versions from the originals in future.[38]

[34] Alex Ball, *Preserving Computer-Aided Design (CAD)* p. 33
[35] Ibid p. 7
[36] MacKenzie Smith, *Final Report for the MIT FACADE Project: October 2006 – August 2009:* https://www.cvaa.be/sites/default/files/projecten/bijlagen/bib_3896_facade_final.pdf accessed 18 June 2018
[37] Ibid pp. 24-25
[38] Ibid p. 25

Drawing upon the results of the FACADE project and other sources, Ball's report suggests a number of best-practice recommendations for the preservation of CAD models[39] that can be loosely summarised as follows:

- Archives should determine a clear rationale for preserving CAD models.
- This rationale should be based on the likely present and future needs of the designated community.
- Guided by this rationale, archives should determine which significant properties of CAD models need to be preserved for future users.
- Where appropriate, CAD models should be normalised (to at least one, but preferably two or three) open-source, vendor-neutral formats to avoid dependence on proprietary native software, while capturing the significant properties targeted for preservation. Ball endorses the STEP, IFC and IGES formats.
- If the preservation rationale demands access to the native CAD model, the archive should consider investing in an emulation platform if native CAD software (and appropriate licencing) is available).
- Normalised derivative files should be validated against the original CAD models.
- Native CAD files should be retained for (only) as long as software is available to read them accurately (Ball notes 'in some circumstances there may be reasons to keep them longer,'[40] for example, to ensure legal compliance).
- Validation metadata should be ingested into the archive alongside the derivative files (and the original CAD models, if they are retained).
- Multi-file CAD models (e.g. Solidworks assemblies) 'should have their part and assembly files archived as a hierarchy of linked packages rather than one large package'.[41] Relationships between parts in the assembly should be expressed as relative paths if possible, and the implied directory structure should be recorded and associated with the top-level assembly file. Elsewhere in the report, Ball suggests that a lightweight derivative of the assembly should be generated from the native CAD model to avoid the need to reconstruct the assembly using lightweight formats.
- Archives should work with donors to ensure that resources and metadata required to understand the CAD model (for example specifications, drawing and layer conventions) are archived alongside the CAD model.
- Archives should encourage the use of documented style conventions for CAD models.
- Finally, 'the wider preservation community should build a business case that underlines the importance of interoperability and preservation' and 'campaign for better support for standard formats in CAD systems among customers, vendors and legislative bodies.'[42]

The normalise-validate approach recommended by Ball may be suitable for small collections where only the geometry or surface appearances of 3D models are important (an example may be the preservation of photogrammetric models for cultural heritage applications). However, this approach will not be capable of preserving many of the advanced model affordances which reveal design intent

[39] Alex Ball, *Preserving Computer-Aided Design (CAD)* pp. 29-30
[40] Ibid p. 29
[41] Ibid p. 29
[42] Ibid p. 30

and which future researchers will likely demand.[43] This is especially true of parametric and/or procedural modelling; 'where the form and its means of generation are profoundly related … losing this information may be tantamount to the loss of the project.'[44]

Moreover, the process of normalising and validating files is (for the present, at least) a manual task requiring native CAD software expertise; given the 'backlog'[45] of paper records (let alone born-digital ones) that many archives have to deal with, this approach is unlikely to scale well for large collections. While Ball suggests that the process of normalisation to open-source formats could be carried out by donors as a condition of ingest, for large donations comprising many CAD files this requirement is likely to deter already busy architectural practitioners from donating in the first place, and is an arguably unrealistic expectation.[46]

Tim Walsh suggests that the process of generating normalised, validated versions targeting significant properties of CAD files 'is perhaps only feasible in repositories that are receiving large numbers of files to the degree that the process can be scripted or otherwise automated.'[47]

Emulation is the process of using one computer's system to imitate the characteristics and performance of another. Often, this means a newer, more powerful computer system imitating an older, less powerful computer system. Emulation is accomplished using a software application, called an emulator, which simulates a different computing environment. For example, the application DOSbox[48] allows the MS-DOS operating system to be run in a window on a current Windows desktop, providing access to legacy software (particularly old PC games) which are no longer compatible with modern versions of Windows. A number of users have successfully installed and run Windows 95 in the DOSbox environment.

As a means of accessing legacy files and software, while not new, emulation is 'perhaps newly viable.'[49] Tim Walsh argues that emulation 'is the only strategy with the potential to provide loss-less rendering of historical CAD data to users.'[50]

Emulation as an access strategy requires that archives collect proprietary software needed to access files. 'The library, archives and museum community will need to cooperate on who has what software,

---

[43] Aliza Leventhal, *Designing the Future Landscape: Digital Architecture, Design & Engineering Assets,* a report on the Architecture, Design and Engineering Summit organized by the Library of Congress, the National Gallery of Art and the Architect of the Capitol on November 16 & 17, 2017 at the Library of Congress: http://digitalpreservation.gov/meetings/ade/ade2017.html accessed 20 June 2018

[44] Schbert, 'Preserving Digital Archives at the Canadian Centre for Architecture' p. 261, quoted in Tim Walsh, *Preservation and Access of Born-Digital Architectural Design Records,* p. 24

[45] Mark A. Greene and Dennis Meissner, "More Product, Less Process: revamping traditional archival processing" in *The American Archivist,* Vol. 68 (Fall/Winter 2005) http://www.archivists.org/prof-education/pre-readings/IMPLP/AA68.2.MeissnerGreene.pdf accessed 26 June 2018

[46] Facade MacKenzie Smith, *Final Report for the MIT FACADE Project* p. 8

[47] Tim Walsh, 'Catching up with the present: Archiving born-digital records of architecture and design', presentation for *Born digital: a symposium exploring Born digital: a symposium exploring digital architectural and built environment records,* Adelaide, 18 – 19 April 2016, p. 15, http://www.unisa.edu.au/born-digital accessed 18 June 2018

[48] https://www.dosbox.com/ accessed 29 June 2018

[49] Tim Walsh, 'Catching up with the present'

[50] Tim Walsh, *Preservation and Access of Born Digital Architectural Design Records,* p. 40

since it is unlikely that each … organisation can maintain copies of all the software they might ever need.'[51]

Emulation as a Service (EaaS)[52] allows access to legacy software remotely through a web browser, and could be used to access CAD/BIM files remotely. EaaS offers 'the ability to provide original-environment experience to access complex digital design objects could offer something previously unimaginable to researchers and practitioners.[53]

While legacy CAD software can be run on virtualised computer platforms, emulation as an access strategy poses practical as well as legal obstacles, since most proprietary CAD software packages contain licencing protection mechanisms.[54]

David S.H. Rosenthal provides a thorough introduction to the potential of emulation and virtualisation (a similar process with some technical differences) as preservation strategies.[55]

Currently, a number of software vendors already offer solutions for viewing proprietary format CAD/BIM files either for free or under special licencing conditions for 'educational' users. Autodesk Inc. offers a number of freely-downloadable file viewers from its website,[56] enabling access to various proprietary Autodesk file formats, as well as a free online (browser-based) file viewer that claims compatibility with over 50 file formats, including proprietary formats from other vendors.[57] In addition, Autodesk also offers its entire suite of current software for free to educational users.[58]

Meanwhile, Robert McNeel & Associates currently offer downloadable evaluation versions of widely-used Rhinoceros 3D modelling software (versions 5 and 6) on their website, which continue to function for viewing and editing files after the end of the 90-day trial period (sans plugins and save/export functionality) thus enabling long-term access to the proprietary .3dm file format. Rhinoceros is also able to import and parse (interpret) a wide variety of other proprietary file formats.[59]

The Industry Foundation Classes (IFC) data model was developed by buildingSMART international, with the intention of expressing BIMs in an open file format, allowing the exchange of information between proprietary software applications.[60] Michelle Lindlar notes that,

> … the IFC data model is the only comprehensive, public, non-proprietary and well-developed data model which supports the full design-to-construction process.[61]

---

[51] MacKenzie Smith, *Final Report for the MIT FACADE Project*
[52] Trevor Owens, 'Emulation as a Service (EaaS) at Yale University Library' in *The Signal,* (August 2014), https://blogs.loc.gov/thesignal/2014/08/emulation-as-a-service-eaas-at-yale-university-library/ accessed 29 June 2018
[53] Aliza Leventhal, *Designing the Future Landscape* p. 14
[54] MacKenzie Smith, 'Curating Architectural 3D CAD Models' in *The International Journal of Digital Curation,* Vol 4, No 1 (2009) p. 103
[55] David S. H. Rosenthal, *Emulation & Virtualization as Preservation Strategies,* Andrew W. Mellon Foundation (2015), https://mellon.org/Rosenthal-Emulation-2015/ accessed 20 July 2018
[56] https://www.autodesk.com/viewers/all-viewers accessed 18 June 2018
[57] https://viewer.autodesk.com/ accessed 18 June 2018
[58] https://www.autodesk.com/education/about-autodesk-education accessed 28 June 2018
[59] https://www.rhino3d.com/download accessed 18 June 2018
[60] Alex Ball, *Preserving Computer-Aided Design (CAD)* p. 16
[61] Michelle Lindlar, 'Building information modelling' p. 6

The *NATSPEC National BIM Guide* recommends that milestone project files are delivered to stakeholders at agreed times over the course of a project (see especially section 10.10 Final BIM Deliverables pp. 19-20), and makes recommendations that BIM models are provided in both native and open-source formats. Following this recommendation, and archiving these files, will see architectural practices well positioned towards ensuring that a minimum standard of records for each project are captured in a 'preservation ready' format (IFC), as well as maintaining original native format files, thus keeping the door open to the possibility of future software emulation.

The FACADE report suggests that such milestone files 'would meet 80% of the users' needs, but the other 20% should be kept for the student or historian who is motivated to browse through that material manually.'[62]

Both Ball and the National Library of Australia's *Digital Preservation Policy* imply that proprietary format files (including native CAD/BIM files) might be deleted when access becomes problematic. However, there are sound arguments for maintaining native file formats in perpetuity, for example:

- Software emulation (including by other collecting institutions)
- Archive users or other collecting institutions with active software licences
- Software vendors and/or patent holders releasing previously proprietary formats as open standards (as occurred with .mp3)

As Megan Philips et al argue, 'while obsolescence is a formidable problem … if a file cannot be opened, it is still a file in one's possession.'[63]

[62] MacKenzie Smith, *Final Report for the MIT FACADE Project* p. 12
[63] Megan Phillips et al, *The NDSA Levels of Digital Preservation* p. 5

## 2.4 Digital architectural records in architectural practice

Over forty years ago, lamenting the loss of paper-based archives, Alan K. Lathrop wrote:

> Architects work for the moment and for the future, usually with little thought for the historical value of their creations. They are businessmen first, artists second, and historians not at all. As with most businessmen, concern for the preservation of their papers beyond their administrative life-span has a very low rank on the scale of priorities.[64]

Architects in practice and archivists in collecting institutions face similar challenges when attempting to preserve digital architectural records, yet they retain or collect records for different reasons. Architects primarily retain records for legal liability; in order to facilitate the recycling of previous work into new projects; as a means of recording work done, so that when a client returns for alterations renovations the firm does not need to start again from scratch; and as a bank of imagery to draw upon for marketing and promotional purposes. Compared to these practical reasons for retaining records, the cultural or historical value of records in the custody of a firm is usually of least concern, while the notion of 'forming a "historical" archive which could be transferred to an archiving institution is rarely talked about as a goal and sometimes explicitly rejected.'[65]

David Peyceré describes architects as the 'first curators' of architectural records.[66]Architectural practices 'not only produce records, but they are their own archivists,'[67] and architectural practitioners 'are already doing part of [the archivists'] job, taking decisions on the fate of each document or file.'[68] In some cases, such decisions are made 'at a very early stage of the record's life, even before the architectural project is itself completed.'[69] While some practices advocate sorting material prior to archiving, [70] others argue that the cost of sorting outweigh the benefits:

> … when you're filing there's a tradeoff between the amount of effort you put in to tidy up before you file it and how easy it is to find stuff afterwards. … I always took the view that you're probably better off not worrying about what's filed, you just file everything and just accept that it's going to be hard to find stuff because you've got to trawl through a whole lot of rubbish. … I never bother to make sure there's only one copy of stuff because the amount of effort and time wasting to do that, the cost of the extra data storage, it's just not worth it.[71]

Traditionally architects have determined their own retention and disposal schedules for records in their care, sometimes on an ad hoc basis. The destruction (or deletion) of files is often regretted

---

[64] Alan K. Lathrop, 'The Archivist and Architectural Records' in *Georgia Archive,* Vol. 5, No. 2 (January 1977) p. 25, https://digitalcommons.kennesaw.edu/georgia_archive/vol5/iss2/4 accessed 7 May 2018

[65] David Peyceré, *The architectural practices as first curators of their archives,* NAi Rotterdam, Hybrid Architectural Archives, 11-12 June 2009, p. 7 conference.nai.nl/mmbase/attachments/525844/1.1_David_Peyceré.pdf accessed 2 May 2018

[66] Ibid

[67] Ibid p. 1

[68] Ibid p. 1

[69] Ibid p. 2

[70] Ibid p. 6

[71] *Interview with Antony McPhee,* 25 July 2016 (Held in University of South Australia Architecture Museum)

later.[72] 'A decision not to preserve is usually a final one for digital materials,' and so the *UNESCO Guidelines for the Preservation of Digital Heritage* recommend a cautious approach to the deletion of digital records. Additionally, the *Guidelines* recommend that organisations

> … decide what materials definitely must be preserved and for how long; what definitely does not need to be preserved; and what should be accepted for interim preservation action while a more definitive selection decision can be made.[73]

**Early born digital architectural records**

> Until about twenty years ago architects remained sceptical of keeping architectural drawings solely in digital, and were consistently making paper copies. This was indeed the cautionary message architectural archivists were sharing with the design community. However, twenty years later, we share the inevitability of being immersed in a digital world, and many of us are already arguing that the digital *is* the original, not the printout record copy.[74]

During the late 1990s, a growing trust in digital technology led to a diminishing ingest of paper records into the hardcopy archives of some Australian architectural practices.[75] While digital files created earlier possessed a redundancy in the form of hard copies, for later projects born digital files were trusted as primary records and their loss is less likely to be mitigated by hardcopy equivalents. Losses have occurred for two main reasons – the inability to open files saved in proprietary software formats, and the failure, or inaccessibility of legacy carrying media.

David Peyceré warns that

> … the first ten or fifteen years of computerised architectural output in many practices may well become inaccessible in the near future – and in some cases this is already so. In 2003 however, three-quarters of European practices surveyed claimed they were able to access the computer files accessed in their firms 10 to 15 years earlier.[76]

Aliza Leventhal characterises this period as a kind of digital 'Wild West': 'an especially experimental time for architects and designers who were testing the limits of design software at their disposal.'[77]

---

[72] *Interview with David Holliday,* 22 June 2016 (Held in University of South Australia Architecture Museum)
[73] National Library of Australia, *Guidelines for the Preservation of Digital Heritage,* UNESCO (2003) p. 72, http://www.unesco.org/new/en/communication-and-information/resources/publications-and-communication-materials/publications/full-list/guidelines-for-the-preservation-of-digital-heritage/ accessed 14 June 2016
[74] Ines Maria Zaluendo, 'Paradigm Shift: curatorial views on collecting and archiving architectural drawings in an evolving born-digital landscape' Paper presented at the Society of American Archivists Conference, Washington D.C. (August 2014) p. 3, https://dash.harvard.edu/bitstream/handle/1/13442962/IZ_Born-Digital%20Architectural%20Drawings_FINAL_2014_A1b.pdf accessed 17 May 2018
[75] *Interview with Glen Collingwood,* 8 March 2016 (Held in University of South Australia Architecture Museum)
[76] David Peyceré, *The architectural practices as first curators of their archives* p. 3
[77] Aliza Leventhal, *Designing the Future Landscape* p. 8

Leventhal also makes a helpful distinction between architectural records of the past and those that are likely to be created into the future:

> The past/present represent the complex "wild west" world of digital design records and the present/future offers a glimpse at potential shifts in designer's workflows, software capabilities, and contract deliverables that might make the digital design records created more manageable and accessible.[78]

**Naming Conventions**

The content of digital files cannot be seen, in many cases, without actually opening a file, and even when files are opened, digital drawings or models lack the iterative 'layered-ness' which may be understood intuitively when viewing paper records.[79] For these reasons (among others), digital architectural records must be identified and organised more rigorously than paper records.[80] However, Paul Minifie notes there is a tendency, in the heat of the moment, for architects to name files in an ad hoc manner:

> You know the classic joke that we still have now is, you do the file overview model, and then you do overview model pm, because you've got to change just a bit of it and you don't have time to properly integrate it and save it and bring everyone else up to the same model, and then there's 'overview model pm final' and then there's 'overview model pm final final' [laughs]… whenever there's forking possibilities of what something might be and you want to keep all the forks alive, it's very hard to structure that in a way that is – can be coherent after the fact.[81]

Antony McPhee agrees:

> … trouble is architects are too creative. When they have to file something, rather than see what it's supposed to be called, they just make something up, so it's really hard to get architects to file stuff. It's why I developed systems that don't rely on having very highly codified methods, it's more the naming structure. If it's slightly wrong, you can still tell what it is.[82]

Suggestions for acceptable folder naming conventions include using plain English descriptive titles; employing a six or eight-digit date code (for example year, month, day); using dated folders; and using folder names to assign context to groups of files, allowing individual file names to be shorter.[83] Furthermore, practitioners should avoid using 'illegal characters'[84] (for example, spaces, question marks, ampersand, angle brackets, etc.).

---

[78] Aliza Leventhal, *Designing the Future Landscape* pp. 19-20
[79] Chris Burns, *Securing and enabling access to knowledge for the future: archiving digital architectural records project report* (November 2016) p. 29, http://aad.unisa.edu.au/siteassets/documents/architecture-musuem/adar-report-201017-public.pdf accessed 28 June 2018
[80] David Peyceré, *The architectural practices as first curators of their archives,* p. 4
[81] *Interview with Paul Minifie,* 14 June 2016 (Held in University of South Australia Architecture Museum)
[82] *Interview with Antony McPhee,* 25 July 2016 (Held in University of South Australia Architecture Museum)
[83] *Interview with Paul Minifie*
[84] https://www.mtu.edu/umc/services/digital/writing/characters-avoid/ accessed 28 June 2018

David Peyceré recommends that the management of in-house architectural archives should draw upon records management principles defined in the ISO 15489 standard, while developing a quality management system based on ISO 9000 principles can help to create 'a solid, reassuring archive structure' and can be used 'as an inspiration without being fully implemented and without certification necessarily being the final goal.'[85]

> Even if every practice has to define its own rules about managing its records, these rules, if applied consistently within each practice, should help us archivists a lot in our task of accepting and managing an archive.[86]

[85] David Peyceré, *The architectural practices as first curators of their archives,* p. 4
[86] Ibid p. 2

## 2.4 Digital architectural records in collecting institutions

### Digital collection policies

A decision to collect digital architectural records should be reflected in the collecting policy of the institution.[87]

While the UNESCO *Guidelines for the Preservation of Digital Heritage* states 'the selection of digital heritage is conceptually the same as selection of non-digital materials,'[88] it is important to recognise that digital records have different associated costs and make different demands on time and resources compared with traditional, paper-based archives. Digital records also raise quite different issues to paper-based archives surrounding authenticity, preservation and access. Therefore, institutions should develop a collection policy for digital objects, or re-write their current collection policy to specifically include digital objects. The UNESCO *Guidelines for the Preservation of Digital Heritage* may offer a sound foundation for the development of a digital collection policy.[89]

Meanwhile, the decision to collect digital architectural records should take into account the capabilities of the institution (in terms of time, funding and expertise) to preserve and, if possible, provide access to records, which are most likely to be in proprietary formats, on an ongoing basis. Institutions should note especially that '[digital] storage is not a one-time cost.'[90]

The importance of paper-based architectural drawings 'lies in their reflection of broad themes of our history or in other words, their "cultural dimensions".'[91] Julie Collins, Susan Collins and Christine Garnaut argue that paper-based architectural records, and in particular drawings,

> are often donated to collecting institutions, particularly libraries that specialise in books, papers and manuscripts but where there is little expertise in architecture or architectural drawings. Serendipitously acquired drawings are then either at risk of not being fully understood or worse still, incorrectly appraised, sometimes resulting in improper destruction.[92]

Archivists appraising digital architectural records should be familiar with the gamut of cultural information that may be encoded in architectural records, both paper and digital. In the case of the latter, archivists should be familiar with the kinds of information that may be revealed when interrogating digital files, especially 3D CAD/BIM models with special affordances, as discussed in section 2.3.

Recent research projects utilising information encoded in native 3D/BIM models were discussed at the Designing the Future Landscape: Digital Architecture, Design & Engineering Assets symposium

---

[87] Julie Collins, Susan Collins and Christine Garnaut, "Behind the Image: assessing architectural drawings as cultural records" in *Archives and Manuscripts,* Vol. 35 No. 2, (November 2007) p. 91

[88] National Library of Australia, *Guidelines for the Preservation of Digital Heritage,* UNESCO (2003) p. 70, http://www.unesco.org/new/en/communication-and-information/resources/publications-and-communication-materials/publications/full-list/guidelines-for-the-preservation-of-digital-heritage/ accessed 14 June 2016

[89] National Library of Australia, *Guidelines for the Preservation of Digital Heritage*

[90] Tim Walsh, *Preservation and Access of Born-Digital Architectural Design Records* p. 30

[91] Julie Collins, Susan Collins and Christine Garnaut, "Behind the Image: assessing architectural drawings as cultural records" p. 93

[92] Julie Collins, Susan Collins and Christine Garnaut, "Behind the Image: assessing architectural drawings as cultural records" p. 90

held in November 2017;[93] meanwhile, the Canadian Centre for Architecture's (CCA) *Archaeology of the Digital* exhibition and book, curated and edited respectively by Greg Lynn, investigates the history of computing in architecture using evidence provided, in part, through the interrogation of digital architectural records.[94]

Digital collection polices should be informed by the needs of the 'designated community' of primary archive users (defined in section 2.2). In the case of digital architectural records, the designated community may be very broad:

> The universe of users of design and construction records is diverse. It includes building owners, architects, preservationists, historians, planners, students, scholars, engineers, lawyers, model builders, horticulturalists, manufacturers, art curators, local communities, and others. Uses include restoration, adaptation and additional construction, historical research, litigation, analysis of sociological factors and patterns of use, publication, exhibition, licencing of designs, educational interpretation, inspiration for new design, the study of drawing techniques, and seismic renovation among others.[95]

Lowell and Nelb argue that 'the most significant project records are the design process and final construction records, and photographs,'[96] especially construction process photographs. Sketches and studies 'document the creative process' while as-built or record drawings 'are the most accurate documentation of a project at its completion.'[97] Meanwhile, construction progress photographs may record changes that have occurred during construction when as-built project record drawings were not compiled accurately by the contractor. Lowell and Nelb also emphasise the importance of retaining the records of unbuilt projects, which 'play an important role in understanding a designers' overall work and may be significant from a design perspective for historical research.'[98]

**Donor guidelines**

Collecting institutions should develop guidelines for donors of digital architectural records. In particular, donor guidelines should encourage early donation of records, and advise a cautious approach to reorganising or otherwise altering digital files before donation (as discussed in section 2.1.2). While designers know 'intuitively which drawings, images or animations best capture design intent,'[99] the possible 'universe of users' of design and construction records may not be immediately obvious, and so including a collection rationale may be beneficial.

Walsh provides the following argument for collecting files as early as possible:

[93] http://digitalpreservation.gov/meetings/ade/ade2017.html accessed 29 June 2018
[94] Greg Lynn (ed.) *Archaeology of the Digital*, Montreal: Canadian Centre for Architecture and Sternberg Press (2013)
[95] Waverly Lowell and Tawny Ryan Nelb, *Architectural Records: managing design construction records,* Chicago, Ill: Society of American Archivists (2006) p. 70
[96] Ibid p. 73
[97] Ibid p. 43
[98] Ibid p. 74
[99] Kristine K. Fallon, *Collecting, Archiving and Exhibiting Digital Design Data Section 2: Archiving Digital Design Data: Practices and Technology,* Art Institute of Chicago, Department of Architecture (2004) http://www.artic.edu/sites/default/files/2A.pdf accessed 28 June 2018

> As with other born-digital materials, the benefits of early and active appraisal and transfer of records from architects and architectural firms cannot be overstated. Selection and transfer of data is most likely to be successful when archivists can consult with records creators. If files are appraised and transferred within a few years of creation, records creators are more likely to remember and share crucial information about their creation, modification, use, and organization, allowing archivists to record more thorough and accurate contextual information for users of the archive. In cases were file naming conventions and CAD style manuals were absent or inconsistently applied, the creators may be able to identify working and final versions of designs and other important qualities of files. Furthermore, early appraisal and transfer leaves open the possibility of managed migration of CAD files to output or standard file formats by taking advantage of the relatively narrow window of time in which a firm will still have a legally licensed installation of the version of the software used to create them, limiting the need for difficult and time-intensive digital archaeological recovery.[100]

Finally, Aliza Leventhal suggests that legacy records 'should be accepted as the complicated sets they are and institutions should focus on supporting them in their original environments, or as close to that as possible.'[101]

**Metadata**

Tim Walsh suggests that the following archival standards designed for the arrangement and description of hardcopy archives may be well suited to describing architectural records:

- *Describing Archives: A Content Standard* (DACS) (US)
- *Rules for Archival Description* (RAD) (Canada)
- *General International Standard Archival Description* (ISAD(G))
- *Encoded Archival Description* (EAD)
- *Encoded Archival Context – Corporate Bodies, Persons, and Families* (EAC—CPF)
- *Library of Congress Subject Headings* (LCSH)
- Getty *Art & Architecture Thesaurus*

For item-level descriptions of CAD files, Walsh recommends *Categories for the Description of Works of Art* (CDWA), developed by the J. Paul Getty Trust, and *Cataloguing Cultural Objects* (CCO), 'which are designed with the unique qualities of visual materials in mind.'[102] The former is also recommended by Kristine Fallon in the *Collecting, Archiving and Exhibiting Design Data* report since it 'provides options for data that closely resembles the "job" or project and individual document level found in design collections.'[103] Walsh suggests,

> Much like audio/visual materials, Word processing documents, and other non-paper records preserved and made accessible by archival repositories, as products and evidence of the day-to-day operations of records creators, CAD and BIM files can be properly intellectually arranged and described

[100] Tim Walsh, *Preservation and Access of Born Digital Architectural Design Records* p. 17.
[101] Aliza Leventhal, *Designing the Future Landscape* p. 10
[102] Tim Walsh, *Preservation and Access of Born Digital Architectural Design Records* p. 27
[103] Waverly Lowell and Tawny Ryan Nelb, *Architectural Records* p. 137

> alongside archival materials of other formats without significantly breaking from traditional practices.[104]

The Metadata Encoding and Transmission Standard (METS) 'is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library.'[105] This standard has been developed by the Network Development and MARC Standards Office of the US Library of Congress, as an initiative of the Digital Library Federation, and is recommended by Walsh as a 'wrapper' for descriptive metadata.[106]

---

[104] Tim Walsh, *Preservation and Access of Born Digital Architectural Design Records* p. 27

[105] Library of Congress, *Metadata Encoding and Transmission Standard,* www.loc.gov/standards/mets accessed 28 June 2018

[106] Tim Walsh, *Preservation and Access of Born Digital Architectural Design Records* p. 27

## 3. Summarised Survey Results

The survey was intended to be anonymous, and did not collect contact details unless volunteered by respondents.

### 3.1 Architectural practice

*Q1 Are you an architect or an archivist?*

*Q2 Are you based in an architectural practice or in a collecting institution?*

Representatives from three architectural practices completed the survey. All three respondents described their position within the practice under the 'other' category – specifically, an information systems coordinator, a librarian, and a knowledge manager.

*Q3 How many staff work in your architectural practice?*

Two of the practices can be considered quite large, with 75 architects/12 other staff and 50 architects/30 other staff. It is not clear whether these practices are spread out over more than one office. The other practice may perhaps be described as medium-sized, with 17 architects/21 other staff. No small practices or sole practitioners completed the survey.

*Q4 What are the most common software products used in your practice (including CAD/BIM as well as word processing, spreadsheet, image manipulation etc. software)?*

All of the practices surveyed used Microsoft Office, the Adobe Creative Suite, and Rhino 3D surface modelling software. Both large practices used Autodesk Revit. The medium-sized practice used Autodesk AutoCAD (alongside Rhino) as its primary CAD software. One of the large practices also used AutoCAD, as well as Autodesk 3ds MAX and SketchUp 3D modelling software. Other software used by the two large practices included Nearmap (an aerial imaging provider), Google Earth Pro, Open Asset (cloud-based digital asset management software), Deltek Vision (for client, project and financial management), Newforma (project file management for architects) MYOB (business and accounting software).

*Q5 To what extent has your practice adopted BIM (Building Information Model) software and workflows?*

Both large practices reported that they used BIM software and workflows for some of their projects. The medium-sized practice reported that they had begun experimenting with BIM.

*Q6 Does your practice use a file-naming convention?*

All respondents reported some kind of file-naming convention. One of the large practices reported that a strict file-naming convention was in place, the other reported that a relaxed file naming convention was in place. Meanwhile the medium-sized practice reported that a strict file-naming convention was in place, but added, 'people don't necessarily follow it correctly.'

*Q7 Is your practice ISO certified (or similar)?*

One of the large practices reported being ISO 9001:2015 certified, while the medium-sized practice reported operating under a Quality Management System reflecting ISO9001:2015 while working towards certification.

*Q8 Approximately how much digital data is stored by your practice in total?*

One of the large practices reported that they stored more than 20TB of digital data; the other two practices reported stored digital data in the 10-20TB range.

*Q9 How does your practice store digital data for completed (archived) projects?*

All practices reported using a server to store data for archived projects. Two practices reported using offline HDD/solid state drives; two reported using optical media; and two reported using magnetic tape. One practice used an external (cloud) storage provider.

*Q10 How does your practice manage long-term storage of data for completed (archived) project data?*

All practices surveyed reported that digital data for archived projects was managed in-house by an IT department to a central repository.

*Q11 How regularly do you retrieve files stored in your digital archive?*

One of the large practices retrieved files from their digital archive weekly, the other on rare occasions, less than monthly. The medium-sized practice retrieved files from their digital archive daily.

*Q12 Do you have a formal policy for migrating files in obsolete formats to current formats?*

None of the practices had a formal policy for migrating files in obsolete formats into current formats.

*Q13 Do you have a formal policy for transferring files from obsolete or volatile data carriers (e.g. magnetic tape, floppy disks, optical disks, flash drives, old hard drives?)*

The medium-sized practice reported they had not only transferred data from older CD and DVD recordable media to their archive server, but had backed up their files on newer disk formats (e.g. DVD recordable dual-layer) affording greater storage capacity.

*Q14 Does your practice retain legacy hardware or software to enable access to obsolete file formats?*

Only the medium-sized practice reported retaining legacy hardware or software; however, the specific example provided was an analogue microfiche reader, so it is not clear whether this practice has retained legacy digital hardware or software.

*Q15 Have you ever donated records (paper or digital) to an archive, library or museum? If not, would you ever doing so?*

One of the larger practices reported that they had donated records to an archive, library or museum. The other large practice had not, but reported that they would consider doing so.

*Q16 If you have donated records to an archive, library or museum (or were to do so in future), what access conditions did (or would) you place on donated records?*

One of the larger practices reported that they had donated records under conditions of full public access. The other large practice stipulated that access would depend on permission from both client and practice.

The medium-sized practice had donated material, however not original work from the practice; nevertheless, it was implied that original work would eventually be donated under conditions of full public access.

*Q17 Have you experienced any of the following challenges retrieving records from your digital archive?*

All of the practices had experienced corrupted files and software obsolescence. The two large practices had both experienced disk failure. One of the large practices and the medium-sized practice had experienced hardware obsolescence.

*Q18 Have any of your past projects involved the translation of designs on paper into CAD drawings, CAD models, or BIMs?*

One of the large practices reported that they had translated drawings on paper into 2D, 3D and BIM formats. The medium-sized practice reported that they had translated drawings on paper into 2D CAD formats.

## 3.2 Collecting institutions

*Q1 Are you an architect or an archivist?*

*Q2 Are you based in an architectural practice or in a collecting institution?*

Representatives from eleven collecting institutions completed the survey. Seven respondents described themselves as archivists. Others described themselves as a curator officer, a curator/architect/archivist, a librarian, a collections manager, and an architect.

*Q20 Select the box that best describes your organisation*

*Q21 How many staff work in your organisation?*

Two respondents described their organisation as falling under state government affiliation (e.g. State Libraries) employing 200+ staff/50+ volunteers and approximately 50 staff/100 volunteers. One respondent described their organisation as Local Government (e.g. council archives) with one paid staff member. One respondent described their organisation as falling under community affiliation (e.g. local history collections) with 1 staff/11 volunteers.

Six respondents described their organisations as universities. One reported 300 staff/50 volunteers – numbers which presumably refer to the university organisations as a whole. Another reported 85 staff, although it's not clear whether this represents an entire university, a school or division within a university, or a department (e.g. a library). Two reported 4 staff/2 volunteers, another reported 3 staff/35 volunteers explicitly within an archives context. One respondent, with an organisation employing ten paid staff, skipped this question.

*Q22 Do you have a staff member who is appointed as a digital archivist?*

One state government organisation and three universities reported having a staff member appointed as a digital archivist.

*Q23 Do you have in-house information technology (IT) staff? If so, how would you describe the level of IT support available?*

Three organisations reported excellent IT support – one state government organisation, one university and one community-level organisation. Five organisations reported good IT support – one state government organisation, one local government organisation, two universities and one unspecified. Two organisations reported limited IT support – both universities.

*Q24 Do members of your staff, or your IT support personnel/provider, have specialised preservation expertise?*

Four organisations reported that their staff had specialised digital preservation expertise (three universities and one state government level organisation). The local government-level organisation reported that their IT support had specialised digital preservation expertise. The other state government organisation reported that *both* staff and IT support had digital preservation expertise.

The remaining four organisations reported no digital preservation expertise.

*Q25 Do you have a digital collection policy?*

Five organisations (four universities and one state government organisation) reported that they are currently developing a digital collection policy. The remainder of respondents do not have a digital collection policy, with one state government organisation reporting that their general collection policy does not distinguish between digital and hardcopy records.

*Q26 Do you provide guidelines for donors of digital files?*

Three respondents indicated they are currently developing guidelines for donors of digital files, including the community level organisation that did not have a digital collection policy. The remainder of respondents reported they do not provide files. Two respondents skipped this question.

*Q27 Does your collection contain any of the following kinds of digital objects?*

Eight respondents reported collecting digital audio files (WAV, MP3, OGG, etc.). Seven respondents reported that they collect 2D digital images (JPEG, PDF, etc.). Seven respondents reported collecting Microsoft Word, Excel, or PowerPoint files (or similar). Six respondents reported collecting digital video files (AVI, MP4, YUV, etc.). Five respondents reported collecting databases (e.g. Microsoft Access).

Only one organisation (a university) reported collecting native 2D and 3D CAD files (various proprietary formats) and native building information model (BIM) files. In the next question, they reported not having a policy for migrating files from proprietary formats into open source or preservation formats.

*Q28 If you collect 3D CAD files (in various proprietary formats) do you have a policy for migrating files from proprietary formats into open source or preservation formats?*

No respondents reported they had a policy for migrating files from proprietary formats into open source or preservation formats.

*Q29 Does your collection contain data stored on any of the following kinds of physical carriers?*

Eight respondents reported that their collection contained optical media (e.g. writable CD, DVD, Blu-Ray disks). Six reported their collection contained hard disk drives. Six reported their collection contained magnetic disks (e.g. floppy disks). Five reported their collection contained magnetic tape. Four reported their collection contained flash media. Two reported their collection contained solid state drives.

*Q30 Do you have a formal policy for transferring files from obsolete or volatile data carriers (e.g. magnetic tape, floppy disks, optical disks, flash drives, old hard drives)?*

Five respondents (three universities, the local government and the community organisation) reported they are developing a formal policy for transferring files from obsolete or volatile data carriers. Three reported they did not have a policy.

*Q31 Some archivists recommend printing out hardcopy (paper) versions of born-digital files. Do you generate hardcopies of digital files as a part of your preservation workflow?*

Seven out of eight organisations reported they did not generate hardcopies of digital files as part of their preservation workflow. Only the local government organisation reported generating hardcopies.

*Q32 A checksum is a value which may be used to verify the integrity of a digital file. Do you use checksums in your digital preservation workflow?*

No respondents reported using checksums in their preservation workflows. Six respondents explicitly answered no.

*Q33 Do you have any suggestions for digital archiving guidelines (as they relate to architectural records) based on your experiences?*

When asked for suggestions for digital archiving guidelines, one of the state government organisations suggested using the PDF file format.

Furthermore, one of the universities noted that they had a limited number of architectural records in their collection, and as such architectural records were 'not a focus' for them.

## 3.3 Discussion

Significantly more individuals based in collecting institutions responded to the survey than individuals based in architectural practices. The relative lack of interest from architectural practitioners may be reflective of the well-documented tendency of architects to be preoccupied by present concerns and short to medium-term project goals rather than their long-term legacy.

Archiving is only one aspect of an architect's work, while for an archivist it is a profession. Perhaps notably, each respondent from architectural practice who completed the survey described themselves as something other than an architect: 'information systems coordinator,' 'record manager' and 'librarian.' Meanwhile the five self-described architects who began the survey did not carry on beyond the orientation questions.

*Architectural Practice*

The three practices surveyed use a range of software; most commonly used are the Microsoft Office and Adobe Creative suites, AutoCAD, Rhino and Revit for 2D CAD, 3D and BIM modelling respectively. The diversity of other, specialised software programs mentioned in this small sample suggests that many more such niche products are used across Australian architectural practice as a whole.

All three respondents use a file naming convention of some kind – a promising result. Likewise, the fact that all three practices manage their data by using an IT department managed to a central repository is promising. Data stored in one place is easier to check for fixity on a regular basis, while data stored on disparate storage media is more likely to be lost, deleted, forgotten or left unchecked for long periods of time.

All three practices appear to use servers for their primary archiving with physical carrying media or external provider backup, but it is not clear whether the practices in question maintain multiple complete copies of their archive.

Unsurprisingly, none of the practices surveyed have a formal policy for migrating files in obsolete formats to current formats. Such a policy would potentially be a time-consuming proposition. However, one practice has actively transferred data from optical media to their archive server as well as backing up files on 'newer disk formats', suggesting the old carrying media has been discarded to save space.

Perhaps most promisingly of all, all practices surveyed expressed their willingness to donate records to a collecting institutions. Two practices suggested they would donate records under conditions of full public access; the other suggested that permission from the practice and client would be necessary to access records.

*Collecting Institutions*

Only one institution reported that their collection contained native format CAD or BIM files. This suggests that few, if any Australian organisations are currently collecting such files actively. In contrast, 2D image, audio, video, document and database files appear to be abundant in Australian collecting institutions.

Five organisations mentioned they are currently developing a policy for transferring files from obsolete or volatile carrying media; all of the organisations appear to have backlogs of physical carrying media in one format or another.

Notably, none of the collecting institutions reported using checksums in their digital preservation workflow, and as reported above, six explicitly answered that they did not use checksums.

Given the unanimous agreement in digital preservation literature (see literature review) on the fundamental importance of generating and maintaining fixity information for all digital objects – it is usually regarded as the first and most important step – this appears an alarming response on face value, especially given that four organisations claimed their staff had digital preservation expertise and/or a digital archivist. However, it may be that bit-level preservation is handled by an IT department or an external (e.g. cloud-based) storage provider, or that the person responding to the survey did not understand the question.

This is possibly a significant finding that deserves further attention, clarification and, if possible, correction.

# A draft national framework for archiving digital architectural records in Australian collecting institutions

## v.1.0 (June 2018)

### Introduction

The emergence of digital technology has had a significant impact on the way in which buildings are designed and constructed. From being regarded initially as a tool to aid design, the computer is now commonly considered to be integral to the design process. The digital environment in which an architectural project is developed involves computer hardware and software in the creation of digital files. The records of the process of designing a building cover a broad spectrum. They include models that explore its potential shape and form, sketches, plans, elevations, sections, renderings and other documents like photographs, emails, letters, faxes, specifications and contracts. Increasingly, these records are produced in digital environments and only exist as digital files.

Currently in Australia a framework does not exist to guide collecting, archiving and preserving these “born digital” architectural records. As a consequence, their future preservation as historical and cultural records and their accessibility to practitioners and researchers is in jeopardy. Architects in practice and archivists in collecting institutions face similar challenges when endeavouring to preserve digital architectural records. The development of this framework was informed by these challenges which are captured in the following four questions:

1. What rationales exist for archiving digital architectural records?
2. Which records should be archived?
3. How should records be archived?
4. How can digital continuity be achieved in rapidly evolving and changing electronic environments?

The framework focuses on digital preservation, and includes:

- A glossary of technical terminology.
- A statement of preservation intent for digital architectural records, describing the different categories of digital architectural records and the significant properties of each category that should be preserved.
- An outline of digital preservation fundamentals, written for a non-specialist audience. The outline stipulates four requirements for successful digital preservation, and concludes with practical advice for individuals or organisations with no or limited IT expertise.
- A set of guidelines for digital preservation, covering several areas of concern: *bit-level preservation; physical carrying media; deciding what to collect; file formats, software and hardware; access; metadata; and retention and disposal (deletion).*
- A table that expands upon the guidelines, providing a rationale for each guideline and relevant resources and further reading that may be useful for the practical implementation of the guidelines.
- A list of references.

A companion document, *A draft national framework for archiving digital architectural records in Australian architectural practice v.1.0 (June 2018),* is targeted at architects in practice. Archivists may benefit from reading this document; likewise, architects may benefit from a deeper understanding of the rationales that underlie the decision-making process of archivists when selecting architectural records for preservation in perpetuity. These two documents complement and reinforce each other.

The preparation of this framework was funded by NATSPEC.

## Glossary

| | |
|---|---|
| BIM | Building Information Model. A 3D *CAD* model that contains embedded information on materials, cost estimation and scheduling. |
| Bit rot | slow and inevitable degradation of a bit stream over time. |
| Bit stream | a sequence of *bits* that stores digital information. |
| Bit | the basic unit of digital information. A bit can have one of only two possible values (on/off), most commonly represented as 0 or 1. |
| CAD | Computer Aided Design. Software used by designers to create precision 2D technical drawings and 3D models. |
| Catastrophic failure | a sudden and total failure from which recovery is impossible. |
| CD | Compact Disk, an *optical physical carrying media* |
| Checksum | a short sequence of characters that acts as a fingerprint of a digital file. A checksum is generated by passing the *bit stream* of a digital file through an algorithm. A checksum may be used to *validate* the *fixity* of a *bit stream.* |
| Cloud computing | the delivery of hosted services over the internet. |
| Corruption | an unintended change in the *bit stream* of a digital file. |
| Dark storage | digital information stored on physical carrying media that is not connected to a computer network. |
| Directory | the structure of a computer filing system; a hierarchy of files and folders. |
| Disk imaging | the process of creating an identical copy of the data stored on a unit of physical carrying media, including empty space and deleted files. |
| DVD | Digital Versatile Disk, an *optical physical carrying media* |
| Emulation | using one computer system to imitate the functions of another. |
| Fixity | (of a bit stream) the state of being unchanging or permanent. |
| Flash memory | digital information stored on a computer chip that can be electrically erased and reprogrammed. |
| Hard disk drive | a *physical carrying media* technology in which large quantities of digital data is stored on a rapidly rotating magnetic platter. |
| Hardware lock | a physical device that renders *software* inoperable when not physically connected to the computer. |

| | |
|---|---|
| Hardware | the physical components of a computer, in contrast to *software*, which are programs or instructions that can be run by hardware. |
| IGES | Initial Graphics Exchange Specification. An open file format for storing and/or CAD data. |
| Ingest | the process of absorbing content into an archive. |
| Installation files | files that allow the installation of a piece of *software* on a computer system. |
| IT | Information Technology |
| JPEG | Joint Photographic Experts Group. A file format for storing *raster graphics* images. |
| Legacy | (of a technology) no longer in current or regular use. |
| Lossless | (of copying) without loss of information. |
| Metadata | data about data; data attached to files stored on a computer system, for e.g. author, date modified, date created. |
| Native format | the format that a piece of software uses by default. |
| OBJ | an *open file format* for storing 3D models. |
| Open format | a file format that is not protected by proprietary patents or other intellectual property restrictions. |
| Optical media | optical *physical carrying media*, in which digital data is stored as markings on disk. The disk is read by a laser while spinning at a high speed. |
| PDF-A | Portable Document Format Archival. An *open file format* for the long-term storage of document content, including text and images. |
| Physical carrying media | the physical media that are used to store information, for example, hard disk drives, magnetic tape or disks, flash media, optical disk (for example CD/DVD recordable) etc. |
| Proprietary format | a file format that is protected by patents or other intellectual property restrictions. |
| Raster graphics | Digital images made up of a rectangular grid of pixels. Compare *vector graphics.* |
| Server | an array of hard disk or solid state drives accessed via a local computer network. |
| Software | programs or instructions that tell a computer how to work, in contrast to physical *hardware* that performs work. |
| Solid state drive | a physical carrying media technology in which large quantities of digital data is stored using *flash memory.* |
| STEP | Standard for the Exchange of Product Model Data. An *open file format* for storing *CAD* data. |

| | |
|---|---|
| SVG | Scalable Vector Graphics. An *open file format* for storing *vector graphics* images. |
| TIFF | Tagged Image File Format. An *open file format* for storing *raster graphics* images. |
| Validate | to check the *fixity* of digital files using *checksums.* |
| Vector graphics | Digital images composed of paths or vectors. Compare *raster graphics.* |

## Statement of preservation intent for born digital architectural records

Digital architectural archives may include any number of file types, including but not limited to: 2D documents, including proprietary word processing, spreadsheet, desktop publishing and presentation files; 2D images, including proprietary format raster and vector image editing files, and open format images; proprietary native format 2D and 3D CAD and BIM models and open format CAD and BIM models (including PDF drawings), email archives, video files, audio files and other proprietary software files. Digital architectural archives will also likely contain proprietary and open source software.

This statement of preservation intent addresses physical carrying media, directory structures and interrelationships, original file metadata, and three common categories of digital objects that are likely to be found in digital architectural collections, which are also most likely to be problematic for maintaining long-term access: 2D images and documents, native file format CAD and BIM models, and CAD and BIM software.

### Physical carrying media, file directories, and metadata

The bulk of born-digital architectural records, not unlike their physical counterparts, are typically organised according to the architectural projects for which they were created. Born digital architectural records are often organised into hierarchical groups of files and folders or *project directories.* Project directories may have evolved organically as the project unfolded, or may be the result of an order imposed on a group of files and folders later.

Sometimes project files may be split across multiple carrying media. Sometimes files may be split across multiple carrying media in a proprietary archive format (i.e. a ZIP format), requiring access to proprietary software to access files in the archive.

Occasionally *working directories* are archived – essentially a copy of the designer’s desktop, containing whatever files they were working on at the time.

Born-digital architectural records that may fall outside project directories, but which may be necessary for understanding the context of records within project directories, may include for example office style manuals, file naming and drawing conventions, document management guidelines and archival retention and disposal schedules and email and/or correspondence archives.

The relationship between files, whether in a project or working directory structure, or when split across carrying media, should be retained. Altering the original directory structure of a project carries with it the likelihood of disrupting links and other interdependencies between files. For this reason, weeding or culling duplicate files may be problematic. Furthermore, the original directory structure – however disordered – may reveal important contextual information about the files themselves, or the working practices of a firm at a particular point in time, which may be of interest to future researchers.

File metadata (e.g. dates created, modified etc.) may be essential to identifying, for example, the most recent version of a file or establishing the chronological sequence of work. It is essential that original file metadata is at least recorded, but is best left unaltered.

Best practice for preserving directory structure and file metadata is to create forensic disk images of physical carrying media, and to preserve disk images in perpetuity as an enduring record of files in their original state, as near as possible to the point of donation. File metadata may also be extracted from files automatically into a spreadsheet using special software.

Generally speaking, the files and directory structure are important, not the physical carrier. However, the physical carrier may be labelled with metadata and should be documented. In some cases, the physical carrier may be significant in itself as an example of a legacy technology, perhaps illustrating the working practices within a particular firm.

*Summary:*

- Relationships between files in a project directory, a working directory or across carrying media – however disordered – should be maintained.
- Preserve original file metadata.
- Avoid weeding or culling duplicate files unless there are sound reasons for doing so.
- Create disk images of physical carrying media and retain them in perpetuity.
- Retain physical carrying media where possible. At minimum, document metadata on physical carrying media before disposal.

**Two dimensional (2D) documents and images**

2D documents may include files in proprietary formats including, for example, Microsoft Word, Excel and PowerPoint files. 2D documents may also include open formats like PDF, including PDF derivatives of drawings created in CAD software.

2D images may include open format raster (e.g. TIFF, JPEG, PNG) and vector images (e.g. SVG), as well as proprietary editing formats created in, for example, Adobe Photoshop, Adobe Illustrator or Adobe InDesign.

In general, the ability to edit 2D images and documents in proprietary formats is not required, and open format derivatives may be created to facilitate long-term preservation and access. However, it is possible that CAD/BIM files may have hidden interactive links to some 2D documents or images, which may be lost if original proprietary format files are discarded – for example, images may be linked to CAD/BIM files as textures applied to 3D geometry. Therefore, reformatting of some image files into different formats without retaining the originals in their original context may be problematic; for example, changing the file extension may cause a link to become severed.

Preservation action to facilitate access could be focused on a curated selection of 2D documents and images that are expected to see the most use; all other files should be preserved in perpetuity at the bit stream level, even if access to proprietary software is temporarily lost.

*Summary:*

- Create open format derivatives of proprietary formats as required, focusing on a small selection of files that will satisfy the needs of most archive users.
- Preserve all proprietary files in their original context in perpetuity at the bit stream level.

**Native file format CAD and BIM models**

Viewing CAD/BIM models in their original software environments is highly desirable, since many aspects of a CAD/BIM model's functionality can only be experienced by opening the model in the original software environment. This is especially true of 3D CAD and BIM models, which have no analogue equivalents, but may also be true of 2D models or drawings. 3D CAD and BIM models contain information that is embedded nowhere else – not even in the completed building.

Nevertheless, viewing CAD/BIM models in their original software environments may not be possible since CAD/BIM models in native formats are usually proprietary and rely on access to proprietary software, which may be dependent on legacy hardware and/or licencing.

However, the preservation intention for CAD/BIM models is that native software files are preserved in perpetuity, even though access to the content of files through proprietary software may be temporarily lost.

It is expected that in the future, access to some legacy file formats may be regained through software emulation. It is also possible that certain proprietary file formats may be released into the public domain when their patents expire (for example the MP3 format, formerly protected by patents, and now an open format). Additionally, other collecting institutions or even individual archive users may possess copies of the proprietary software needed to access the content of files, even if the collecting institution preserving the files does not.

Since access to the content of native CAD/BIM models is precarious, CAD/BIM models should be converted (reformatted) into open-format derivatives on a case-by-case basis. While each architectural project may contain a large number of CAD/BIM files, access to a small number of files from each project (for example the final design intent model, or the final as-built model) is expected to serve the needs of most users. Preservation action to facilitate access could be focused on a curated selection of CAD/BIM files that are expected to see the most use; other files should be preserved in perpetuity at the bit stream level. All native proprietary files should be maintained in perpetuity, even if access to native software is temporarily lost (loss of access to proprietary software and file formats should always be considered temporary).

Where open source derivatives already exist side-by-side with native CAD files in a project directory, both should be retained in perpetuity.

*Summary:*

- Access to proprietary CAD/BIM files in their native software environments is desirable but not always possible.
- Preserve all native CAD files in perpetuity at the bit stream level.
- Create open format derivatives for a small selection of native CAD files that will satisfy the needs of most archive users.
- Where open source derivatives already exist side-by-side with native CAD files and a project archive, both formats should be retained in perpetuity.

### CAD and BIM software

Because collecting digital architectural records means collecting files in proprietary formats which are in many cases dependent on proprietary software, preserving digital architectural records means preserving the proprietary software that was used to create them. Software includes for example installation files on physical media, manuals, boxes and documentation, certificates of authenticity, licence keys and hardware locks.

All CAD/BIM software should be preserved in perpetuity, even if not supported by present hardware. As discussed above, it is expected that in the future access to some legacy software may become possible through emulation.

Physical documentation and manuals may be retained in hardcopy form, or digitised, or both. The content of documentation and/or manuals is important, not the physical format.

Certificates of authenticity should be retained in physical form.

Hardware locks are essential to operating some software and must be retained in physical form.

When installation files exist on volatile carrying media (e.g. floppy disks) they may need to be ingested into the central archival repository to receive bit stream-level preservation. Files carried on commercially pressed CD-ROM disks may be ingested into the central archival repository, however such media will last indefinitely if cared for, and so may be retained as well or instead.

*Summary:*

- Software includes installation files on physical media, manuals, boxes and documentation, certificates of authenticity, licence keys and hardware locks.
- Preserve all software installation files in perpetuity at the bit stream level.
- Retain a copy of all documentation, manuals and software boxes – either in physical or digitised form, or both.
- Retain certificates of authenticity in physical form.
- Retain hardware locks in physical form.
- Ingest all installation files on volatile physical carrying media into the central archival repository for bit stream level preservation in perpetuity.
- Retain commercially pressed CD-ROM disks, or ingest file contents into the central archival repository for bit stream level preservation in perpetuity, or both.

## Digital preservation fundamentals

All digital files are composed of *bits* – the smallest possible units of digital information. Bits are usually represented as 0s and 1s. A sequence of bits is called a *bit stream*. A bit stream spells out information in a *binary code* that is understandable by computers.

Digital files may be copied an infinite number of times without degradation – they may be said to be *lossless.* Analogue media, on the other hand, will degrade with each *generation* of copying (for example a photocopy will be of a lower quality than the original; a photocopy of a photocopy will be of a lower quality still). Note that some digital file formats have lossy *encoding* of their *content* (e.g. JPEG image) where digital information is discarded at the point of encoding (i.e. the act of "saving as" a JPEG discards information forever). Nevertheless, the bit stream of a JPEG image, or any other file, can be copied an infinite number of times without *further* degradation.

The media on which a digital bit stream is stored is known as *physical carrying media* – 'physical' because carrying media exists physically, in the real world; 'carrying' because it stores or 'carries' digital *bits.* Examples of physical carrying media include, for example, hard disk drives, solid state drives, magnetic media (e.g. floppy disks, magnetic tape) optical media (e.g. CD/DVD/Blu-ray recordable disks) and flash media (e.g. digital camera cards, USB thumb drives).

A *server* is usually an array of hard disk or solid state drives, accessed via a local computer network. *Cloud storage* is usually a networked array of hard disks or solid state drives, operated by a third party and accessed via the internet.

Depending on the physical carrying media, bits may be stored magnetically, or as electrical charges, or as physical indentations on the surface of a disk, or holes in a punched card, among other methods.

All currently available physical carrying media have comparatively short reliable lifespans compared with hardcopy (paper-based) records, typically within the range of 2-5 years.

All currently available physical carrying media are susceptible to unpredictable *catastrophic failure*, when access to all stored data may be suddenly lost, without warning.

All currently available physical carrying media is susceptible to slow degradation, where individual or multiple bits may spontaneously change over time – for example a 0 may change into a 1, or a 1 may change into a 0, or a bit may be omitted, or an extra bit may be added, or an entire string of bits may change.

Once a bit in a digital file has changed, the file is said to be *corrupted.* Once a digital file has corrupted, it is practically impossible to repair.

While corrupted files cannot be repaired, the lossless nature of digital information makes it easy to prepare identical copies of digital files.

Identical copies should be kept in different *physical locations*, that is, not together on the same physical carrying media, which may fail catastrophically.

If a file becomes corrupted, it is easy to replace it with an identical copy.

The integrity or *fixity* digital files must be monitored so that corruption may be detected and corrupted files replaced with identical copies.

A *checksum* is a short sequence of characters that acts as a fingerprint of a digital file. A checksum is *generated* by passing the bit stream of the digital file through an *algorithm*. If the bit stream becomes corrupted, *even by a single bit,* a different checksum will be generated when the bit stream is passed through the checksum algorithm.

Periodically comparing a digital file to its corresponding checksum will reveal whether or not a digital file has become corrupted.

Checksums can be generated automatically for large numbers of digital files.

Large numbers of digital files can be automatically checked against checksums generated earlier.

**Therefore, preservation of digital bit streams requires:**

1. Maintaining at least *two identical copies* of all digital files in *different physical locations*.

2. Generating and keeping *checksums for all copies* of all digital files.

3. *Checking* the fixity of all copies of digital files *regularly* against checksums.

4. *Replacing corrupted files* when detected with identical copies.

*If these four essential bit stream-level digital preservation actions are not observed, all other digital preservation actions are futile.*

Checksums can be easily generated and checked for the contents of entire *directories* (a hierarchy of files in folders) with freely available, user-friendly utilities such as TeraCopy: http://www.codesector.com/teracopy, accessed 23 June 2018.

If the above preservation actions cannot be applied, it is suggested that archivists with limited information technology expertise do the following:

1. Burn at least two copies of all files (preferably with checksums) to M-DISC DVD-R recordable media.

2. Burn at least two copies of all files (preferably with checksums) to regular DVD-R recordable media, using at least two different disk brands (e.g. TDK and Verbatim).

3. Label and date all disks on the central (usually transparent) hub.

4. Store one copy on each media in a separate place (i.e. one copy on M-DISC and one copy on regular DVD-R in one place, the other copies in another place).

5. Complement each disk with an identical copy of that disk every five years. Do not discard the original disks.

6. Label and date complementary copies.

*Note that these steps will not guarantee the preservation of digital data, but they will increase the likelihood that digital data will survive until it falls into the custody of a preservation repository.*

See also David Farneth, *Selecting, acquiring, managing, and preserving born digital collections… in 5 minutes!* https://www.ifla.org/files/assets/rare-books-and-manuscripts/Lyon-2014-presentations/wlic_2014_lyon_session_84_farneth.pdf accessed 27 June 2018.

**Key Guidelines**

*1. Bit-level preservation*

1.1 All archivists working with digital records should have a basic understanding of fundamental bit-level digital preservation principles.

1.2 Digital records should be stored in a secure central archival repository.

1.3 At minimum, two complete copies of the central archival repository should be maintained. One copy should be maintained offsite.
Best practice: maintain at least three complete copies of the central archival repository. One copy should be maintained in a geographically distant location.

1.4 An external storage provider may be used to maintain an offsite copy of the central archival repository.
Best practice: external storage providers should be 'trustworthy' (ISO 16363 / TDR certified or similar) and provide bit stream-level preservation of digital files.

1.5 All files in the central repository should have associated fixity information. All files in the repository should be checked against fixity information on a regular basis. Corrupted files should be replaced with copies prepared at the time of donation.

1.6 Files in backup copies of the repository should also be checked against fixity on a regular basis. Corrupted files should be replaced from the main repository or another backup repository.

*2. Physical carrying media*

2.1 Data that only exists on disparate carrying media (e.g. floppy disks, optical disks, CD, DVD and Blu-ray recordable disks, old hard disk drives, flash media etc.) should be ingested into the central archival repository.

2.2 Use a workstation that is not connected to the central repository to run initial virus checks and perform disk imaging.

2.3 Where carrying media is labelled, consider documenting carrying media photographically.

2.4 Data ingested into the archive from disparate carrying media should be copied in such a way that the original directory structure and metadata are retained intact.

2.5 Exercise care when copying files from carrying media.
Minimum practice: validate files on carrying media against copies ingested into the repository.
Best practice: generate disk images of physical carrying media, and ingest disk images into the digital archive alongside copies of validated files.

2.6 Exercise caution when disposing of original physical carrying media.
Minimum practice: retain corrupted physical media for future forensic recovery.
Best practice: retain all physical carrying media.

*3. Deciding what to collect*

3.1 A decision to collect digital architectural records should reflect the collection policy of the institution.

3.2 Institutions should develop a collection policy for digital objects, or re-write their current collection policy to include digital objects.

3.3 The decision to collect digital architectural records should take into account the capabilities of the institution (in terms of time, funding and expertise) to preserve and, if possible, facilitate access to records (which are likely to be in proprietary formats) on an ongoing basis.

3.4 Institutions should develop donor guidelines for born-digital architectural records. In particular, donor guidelines should encourage early donation of records, and advise a cautious approach to reorganising or otherwise altering digital files before donation.
Donor guidelines may also stipulate that certain CAD/BIM files can be donated in open formats alongside the original, native format versions.

3.5 Archivists appraising digital architectural records should be familiar with the gamut of cultural information that may be encoded in architectural records, including in drawings in traditional (paper) and in digital (2D and 3D) formats.

3.6 Archivists should be aware of the affordances of native (proprietary format) 3D CAD/BIM models for encoding information (including design intent) that open format derivative files and/or 2D drawings (paper or digital) are incapable of replicating.

3.7 If it is decided to collect digital records from a particular architectural project, collecting institutions should consider collecting the following minimum records: 1) milestone project deliverables including as-built and design intent contract documentation and specifications; 2) a selection of design development records (including e.g. sketches and studies); 3) photographs, especially construction progress photographs.

*4. File formats, software and hardware*

4.1 If a decision is made to collect CAD files or BIM models associated with a particular project, native CAD/BIM files should be collected and archived permanently alongside any extant open format derivative files, even if ongoing access through native proprietary software cannot be supported.

4.2 When collecting 2D and 3D CAD and BIM models, archives should, if possible, collect software that is necessary to open them – even when the software is not supported by current hardware.
Collect installation media, manuals and documentation, licence keys and hardware locks.
Ingest copies of installation files into the central archival repository. Where necessary, transfer installation files from legacy carrying media (e.g. floppy disks) and ingest into the central archival repository. While commercially pressed CD-ROM disks may last indefinitely if cared for, consider making a backup copy and ingest into the central repository.
Share details of proprietary software holdings with other collecting institutions, and cooperate with other institutions to establish a network of legacy software users.

4.3 Collect office style manuals, file naming and drawing conventions, document management and archival retention and disposal schedules, along with other relevant practice-wide documentation.

4.4 Collect functioning legacy hardware, especially physical carrying media drives (e.g. tape drives, floppy and zip disk drives etc.).
Share details of functional legacy hardware holdings with other collecting institutions, and cooperate with other collecting institutions to establish a network of legacy hardware users.
Consider donating unwanted functioning legacy hardware to a collecting institution with a legacy hardware focus.

*5. Access*

5.1 When access to CAD/BIM files through native software is not possible, open format derivatives of milestone CAD/BIM files should be created and ingested into the central archival repository alongside native format files.

5.2 CAD/BIM models should be converted (reformatted), where necessary, into open-format derivatives on a case-by-case basis. Preservation action to facilitate access should be focused on a curated selection of CAD/BIM files that are expected to see the most use; other files should be preserved in perpetuity at the bit stream level.

5.3 Consider a 'reformat on demand' approach to proprietary CAD/BIM files (analogous to 'digitisation on demand').

5.4 Appropriate file formats for creating open derivatives may include for example:

- For word processing documents, PDF/A.
- For spreadsheets, CSV.
- For 2D CAD drawings and/or vector graphics, PDF/A and SVG.
- For bitmap images, TIFF.
- For 3D CAD formats, STEP, IGES and OBJ as appropriate.
- For BIM data, Industry Foundation Classes (IFC) for BIM models and PDF/A for 2D drawings.

*6. Metadata*

6.1 Collect as much contextual metadata as possible for each project or collection, as appropriate, at the point of donation.
Contextual metadata may include, for example, name of architect, name of project, client, contractors, consultants, location, address, GPS coordinates, year designed, year completed, donated by, date donated, staff member received by, access conditions, etc.
Consider supplying donors with a pro forma to enable them to record this information.

6.2 File metadata (e.g. date created, date modified, date accessed, author etc.) can and should be generated automatically.

6.3 As a minimum, consider archiving within each project or collection folder:

a) A text document containing an outline of folder contents and documenting contextual metadata.
b) A spreadsheet containing a list of files and associated file metadata.

*7. Retention and disposal (deletion)*

7.1 Avoid weeding duplicate files unless absolutely necessary (e.g. duplicate copies of entire directory structures, or large files (e.g. 1 GB or larger). If it is necessary to weed duplicate files, proceed with caution.

7.2 Exercise caution when disposing of digital records.
Err on the side of retaining more rather than less.
"If in doubt, don't throw it out!"

## Key Guidelines: Rationale, Resources and Further Reading

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| *1. Bit-level preservation* | | |
| 1.1 All archivists working with digital records should have a basic understanding of fundamental bit-level digital preservation principles. | The 2018 survey of archival professionals undertaken for the NATPSEC funded Archiving Digital Architectural Records project suggests that basic digital preservation principles may not be well understood in the archival community.<br><br>Bit-level preservation is the foundation for all digital preservation activities; if the bit-stream fixity of digital objects cannot be assured, all other digital preservation actions are futile.<br><br>All commercially available digital carrying media are subject to both slow degradations over time ('bit rot') and the possibility of sudden, catastrophic failure.<br><br>Even if not directly responsible for bit-level preservation infrastructure, cultivating an understanding of bit-level preservation principles will enable archivists to confidently discuss their digital preservation requirements with IT support personnel or external digital storage providers. | Brian Lavoie, *The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition).* The OAIS reference model has become the universally accepted *lingua franca* of digital preservation. Lavoie's report provides an accessible introduction to OAIS terminology and concepts: http://dx.doi.org/10.7207/twr14-02 accessed 24 June 2018<br><br>The NDSA *Levels of Digital Preservation* offers a set of plain-language, rubric-based guidelines for digital preservation: www.digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf accessed 14 June 2018<br><br>The *Digital Preservation Capability Maturity Model (DPCMM)* provides digital preservation performance metrics benchmarked against ISO 14721 conformance: https://www.statearchivists.org/resource-center/resource-library/digital-preservation-capability-maturity-model-dpcmm/ accessed 14 June 2018<br><br>David S. H. Rosenthal, Thomas Robertson, Tom Lipkisi, Vicky Reich and Seth Morabito, "Requirements for Digital Preservation Systems: A Bottom-Up Approach" in *D-Lib Magazine* (Vol. 11 No. 11, November 2005) provides an in-depth discussion of the challenges of digital preservation: www.dlib.org/dlib/november05/rosenthal/11rosenthal.html accessed 14 June 2018<br><br>Digital Preservation Coalition, *Digital Preservation Handbook (2nd Edition)* (2015) https://www.dpconline.org/handbook accessed 27 June 2018<br><br>The National Archives (UK) *Preserving Digital Records: Guidance* http://www.nationalarchives.gov.uk/information- |
| 1.2 Digital records should be stored in a secure central archival repository. | Centralising files in one location streamlines the process of checking the fixity of digital files (see below).<br><br>A central repository could be e.g.:<br>• A networked server (for large organisations)<br>• A hard disk or solid state drive (for small organisations)<br>• An external storage provider | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| 1.3 At minimum two complete copies of the central archival repository should be maintained. One copy should be maintained offsite.<br><br>Best practice: maintain at least three complete copies of the central archival repository. One copy should be maintained in a geographically distant location. | Complete copies of the repository are maintained so that corrupted files may be replaced with identical copies.<br><br>A magnetic tape backup that is regenerated on a regular basis from the central repository could serve as one copy, but is still subject to degradation over time and should not be trusted as long-term 'dark storage'. | management/manage-information/preserving-digital-records/guidance/ accessed 27 June 2018<br><br>David Farneth, *Selecting, acquiring, managing, and preserving born digital collections… in 5 minutes!* https://www.ifla.org/files/assets/rare-books-and-manuscripts/Lyon-2014-presentations/wlic_2014_lyon_session_84_farneth.pdf accessed 27 June 2018 |
| 1.4 An external storage provider may be used to maintain an offsite copy of the central archival repository.<br><br>Best practice: external storage providers should be 'trustworthy' (ISO 16363 / TDR certified or similar) and provide bit stream-level preservation of digital files. | External storage providers may not necessarily provide secure, bit-level preservation of digital files.<br><br>Using an ISO 16363 certified provider increases the likelihood that the original bit stream of digital files submitted for long-term storage will be preserved. | ISO International Standard 16363 describes criteria for the audit and certification of trustworthy digital repositories. |
| 1.5 All files in the central repository should have associated fixity information. All files in the repository should be checked against fixity information on a regular basis. Corrupted files should be replaced with copies prepared at the time of donation. | Checking against previously generated fixity information is the only efficient and economical way of monitoring the integrity of files. Note that fixity information can be generated automatically for large numbers of files. Likewise, large numbers of files can be checked against fixity information automatically. Once corruption has occurred, files can only be recovered by replacing them with backup copies. | For examples of software able to generate checksums and validate files see Julianna Barrera-Gomez and Ricky Erway, *Walk This Way: detailed steps for transferring born-digital content from media you can read in-* |

| Key Guidelines | Rationale | Resources and further reading |
| --- | --- | --- |
| 1.6 Files in backup copies of the repository should also be checked against fixity on a regular basis. Corrupted files should be replaced from the main repository or another backup repository. | Backup copies of the repository will degrade over time just as the main repository will degrade over time. | *house* (Dublin, Ohio: OCLC Online Computer Library Centre, Inc., 2013): http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf accessed 14 June 2018<br><br>TeraCopy is a 'freemium' file copying utility (free for non-commercial use) that enables easy generation and checking of fixity information for any directory structure: http://www.codesector.com/teracopy accessed 22 June 2018 |
| *2. Physical carrying media* | | |
| 2.1 Data that only exists on disparate carrying media (e.g. floppy disks, optical disks, CD, DVD and Blu-ray recordable disks, old hard disk drives, flash media etc.) should be ingested into the central archival repository. | Centralising files in one location streamlines the process of checking the fixity of digital files. Checking the fixity of files stored on disparate carrying media is laborious and time consuming. CD/DVD recordable media is especially vulnerable to damage when handled, resulting in possible data loss. | For practical advice on transferring files stored on disparate physical carrying media, including the use of write blockers and the generation of disk images, as well as links to software, see *Walk This Way: detailed steps for transferring born-digital content from media you can read in-house,* http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf accessed 14 June 2018<br><br>The Digital Preservation team at the National Library of Australia (NLA) has developed an online, searchable database of carrying media, called Mediapedia, designed to assist with the identification of various physical carrier formats. See: https://mediapedia.nla.gov.au/home.php accessed 14 June 2018<br><br>Tim Walsh, 'Catching up with the present: Archiving born-digital records of architecture and design' presentation at *Born digital: a symposium exploring Born digital: a symposium exploring digital architectural and built environment records,* Adelaide, 18 – 19 April 2016: http://www.unisa.edu.au/born-digital accessed 18 June 2018 |
| 2.2 Use a workstation that is not connected to the central repository to run initial virus checks and perform disk imaging. | Isolating donated digital content from the central archival repository increases the likelihood that viruses will be detected and eliminated before they are able to compromise files held in the central archival repository. | |
| 2.3 Where carrying media is labelled, consider documenting carrying media photographically. | Information recorded on carrying media may not necessarily be recorded within the content of the digital information itself.<br><br>Photography is able to capture the occasional ambiguity of hand-written labels, whereas transcription is liable to human error. | |
| 2.4 Data ingested into the archive from disparate carrying media should be copied in such a way that the original directory structure and metadata are retained intact. | Altering the original directory structure of a project carries with it the likelihood of disrupting links and other interdependencies between files. Furthermore, the integrity of the original directory structure – however disordered – may reveal important contextual information about the files | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| | themselves or the working practices of a firm at a particular point in time, which may be of interest to future historians.<br>Likewise, metadata (e.g. dates created, modified etc.) may be essential to identifying, for example, the most recent version of a file or establishing the chronological sequence of work. | |
| 2.5 Exercise care when copying files from carrying media.<br><br>Minimum practice: validate files on carrying media against copies ingested into the repository.<br><br>Best practice: generate disk images of physical carrying media, and ingest disk images into the digital archive alongside copies of validated files. | Files may be incorrectly or incompletely transferred when relying on the built-in copy/paste functionality of a computer's operating system.<br><br>Disk imaging may allow data on corrupted physical media to be partially recovered.<br><br>Disk images may act as a valuable enduring record of files in their original state, as near as possible to the point of donation, before any curatorial action has taken place. | |
| 2.6 Exercise caution when disposing of original physical carrying media.<br><br>Minimum practice: retain corrupted physical media for future forensic recovery.<br><br>Best practice: retain all physical carrying media. | Files may be incorrectly or incompletely transferred when relying on the built-in copy/paste functionality of a computer's operating system.<br><br>Improvements in technology may render physical media previously thought to be corrupted beyond repair accessible once again. | |

| Key Guidelines | Rationale | Resources and further reading |
| --- | --- | --- |
| | The physical carrier may be significant in itself as an example of a legacy technology, perhaps illustrating the working practices within a particular firm. | |
| *3. Deciding what to collect* | | |
| 3.1 A decision to collect digital architectural records should reflect the collection policy of the institution.<br><br>3.2 Institutions should develop a collection policy for digital objects, or re-write their current collection policy to include digital objects. | While the UNESCO *Guidelines for the Preservation of Digital Heritage* states 'the selection of digital heritage is conceptually the same as selection of non-digital materials,' (UNESCO p. 70) it is important to recognise that digital records have different associated costs and make different demands on time and resources compared with traditional, paper-based archives. Digital records also raise quite different issues to paper-based archives surrounding authenticity, preservation and access. | National Library of Australia, *Guidelines for the Preservation of Digital Heritage,* UNESCO (2003) may offer a sound foundation for the development of a digital collection policy: http://www.unesco.org/new/en/communication-and-information/resources/publications-and-communication-materials/publications/full-list/guidelines-for-the-preservation-of-digital-heritage/ accessed 14 June 2016<br><br>Examples of digital preservation policies include:<br><br>National Library of Australia (NLA), *Digital Preservation Policy 4th Edition (2013),* https://www.nla.gov.au/policy-and-planning/digital-preservation-policy accessed 28 June 2018<br><br>The NLA Digital Preservation Policy is informed by their *Statements of Preservation Intent:* https://www.nla.gov.au/content/statements-of-preservation-intent accessed 28 June 2018<br><br>Archives New Zealand, *Digital Preservation Strategy:* http://archives.govt.nz/advice/government-digital-archive-programme/digital-preservation-strategy/digital-preservation-strat accessed 28 June 2018 |
| 3.3 The decision to collect digital architectural records should take into account the capabilities of the institution (in terms of time, funding | Collecting digital architectural records (in most cases) means collecting proprietary 2D CAD and 3D CAD/BIM files in native file formats, which entails a combination of one of two preservation pathways to facilitate | Aliza Leventhal, *Designing the Future Landscape: Digital Architecture, Design & Engineering Assets,* a report on the Architecture, Design and Engineering Summit organized by the Library of Congress, the National Gallery of Art and the Architect of the Capitol on November 16 & 17, |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| and expertise) to preserve and, if possible, facilitate access to records (which are likely to be in proprietary formats) on an ongoing basis. | access: migrating proprietary files into standard formats or emulating proprietary software so that files may be opened in their native environments. Note that emulation is an emerging technology, but may become prevalent in the near future, offering scale-able access to large quantities of records. | 2017 at the Library of Congress: http://digitalpreservation.gov/meetings/ade/ade2017.html accessed 20 June 2018<br><br>David S. H. Rosenthal, *Emulation & Virtualization as Preservation Strategies,* Andrew W. Mellon Foundation (2015): https://mellon.org/Rosenthal-Emulation-2015/ accessed 20 July 2018 |
| 3.4 Institutions should develop donor guidelines for born-digital architectural records. In particular, donor guidelines should encourage early donation of records, and advise cautious approach to reorganising or otherwise altering digital files before donation. Donor guidelines may also stipulate that certain CAD/BIM files can be donated in open formats alongside the original, native format versions. | 'Selection and transfer of data is most likely to be successful when archivists can consult with records creators' (Walsh 2015 p. 17). If files are transferred early, 'records creators are more likely to remember and share crucial information about their creation, modification, use, and organisation, allowing archivists to record more thorough and accurate contextual information for users of the archive' (p. 17).<br><br>See rationale 2.4 regarding details on file directory and metadata integrity.<br><br>See rationale 3.6 regarding collection of open format as well as native format CAD/BIM files. | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| 3.5 Archivists appraising digital architectural records should be familiar with the gamut of cultural information that may be encoded in architectural records, including in drawings in traditional (paper) and in digital (2D and 3D) formats. | The importance of architectural drawings 'lies in their reflection of broad themes of our history or in other words, their 'cultural dimensions' (Collins, Collins and Garnaut 2006 p. 93).<br><br>A familiarity with the kinds of cultural information that may be encoded in architectural records is necessary to make informed decisions about which records should be kept. | Julie Collins, Susan Collins and Christine Garnaut, "Behind the Image: assessing architectural drawings as cultural records" (in *Archives and Manuscripts,* Vol. 35 No. 2, November 2007, pp. 87-107) discusses the kinds of information that may be embodied in traditional (paper) architectural drawings, and includes a ready reference chart for the assessment of traditional (paper) architectural drawings for cultural dimensions.<br><br>Waverly Lowell and Tawny Ryan Nelb's *Architectural Records: managing design construction records* (Chicago: Society of American Archivists, 2006) discusses the appraisal of paper-based architectural records for retention or disposal (pp. 69-85) and provides a table (pp. 84-85) suggesting retention, disposal and reformatting actions for a variety of record types. While aimed primarily at archivists working with analogue media, it provides an indication of the kinds of records that collecting institutions may consider worthy of long-term preservation. |
| 3.6 Archivists should be aware of the affordances of native (proprietary format) 3D CAD/BIM models for encoding information (including design intent) that open format derivative files and/or 2D drawings (paper or digital) are incapable of replicating. | Native CAD/BIM files are the "originals" and contain information that may be lost when converting them into open format derivatives, for example:<br>• Construction history that may be rolled back and forward like an interactive "undo/redo" feature. The history of modelling actions may be rolled back to an earlier point in the model's development, changes implemented, and the model "rolled forward" again, automatically updating later stages of construction with the results of the changes.<br>• Parametric modelling, in which 'aspects of the design are given a variable value instead of a fixed one, in order to make them easier to adjust and reuse in different contexts' with 'constraints [that] control how designs should | Recent research projects utilising the information encoded in native 3D/BIM models are mentioned in Aliza Leventhal, *Designing the Future Landscape: Digital Architecture, Design & Engineering Assets,* http://digitalpreservation.gov/meetings/ade/ade2017.html accessed 20 June 2018<br><br>The Canadian Centre for Architecture's (CCA) *Archaeology of the Digital* exhibition and book (Montreal: Canadian Centre for Architecture and Sternberg Press, 2013), curated and edited respectively by Greg Lynn, investigates the history of computing in architecture using evidence provided, in part, through the interrogation of digital architectural records: https://www.cca.qc.ca/en/events/3333/archaeology-of-the-digital accessed 18 June 2018 |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| | be adjusted in the light of changed variables' (Ball p. 33).<br>• Features, which are the parametric building blocks of a model within some solid modelling applications (e.g. Solidworks and Autodesk Inventor) and in many cases have real-life manufacturing analogues – for example, "draft angle" (taper) added to a moulded part to allow it to be removed from a mould, a "fillet" (curved blended surface) to reduce stress where two surfaces meet, or a solid body hollowed out to create a "shell".<br><br>Retaining native CAD/BIM files, even when access to proprietary software is lost, leaves the door open to future possibilities that may enable regained access, for example:<br>• Software emulation (including by other collecting institutions)<br>• Archive users or other collecting institutions with active software licences<br>• Software vendors and/or patent holders releasing previously proprietary formats as open standards (as occurred with .mp3) | |
| 3.7 If it is decided to collect digital records from a particular architectural project, collecting institutions should consider collecting the following minimum records: 1) milestone project deliverables (see further reading column), including as-built and | Lowell & Nelb (2006 p. 73) state that 'the most significant project records are the design process and final construction records, and photographs. These records are required for understanding the development of the program'<br><br>Sketches and studies 'document the creative process' while as-built or record drawings 'are the most | The *NATSPEC National BIM Guide* provides guidelines on the types of files that should comprise milestone digital project deliverables (Construction and Information Systems Limited 2016): https://bim.natspec.org/documents/natspec-national-bim-guide accessed 14 June 2018 |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| design intent contract documentation and specifications; 2) a selection of design development records (including for e.g. sketches and studies); 3) photographs, especially construction progress photographs. | accurate documentation of the project at its completion' (Lowell & Nelb 2006 p. 43).<br><br>Meanwhile, construction progress photographs may record changes that have occurred during construction when as-built project record drawings were not created by the contractor. | |
| *4. File formats, software and hardware* | | |
| 4.1 If a decision is made to collect CAD files or BIM models associated with a particular project, native CAD/BIM files should be collected and archived permanently alongside any extant open format derivative files, even if ongoing access through native proprietary software cannot be supported. | See rationale for 3.6. | |
| 4.2 When collecting 2D and 3D CAD and BIM models, archives should, if possible, collect software that is necessary to open them – even when the software is not supported by current hardware.<br><br>Collect installation media, manuals and documentation, licence keys and hardware locks.<br><br>Ingest copies of installation files into the central archival repository. Where necessary, transfer installation files from legacy | Collecting digital architectural records means collecting files in proprietary formats which are in many cases dependent on proprietary software. Therefore, collecting digital architectural records means collecting proprietary software.<br><br>'The library, archives and museum community will then need to cooperate on who has what software, since it is unlikely that each archiving organization can maintain copies of all of the software they might ever need' (Smith 2009 p. 28). | Autodesk Inc. offers a number of freely-downloadable file viewers from its website, enabling access to various proprietary Autodesk file formats: https://www.autodesk.com/viewers/all-viewers accessed 18 June 2018<br><br>Autodesk Inc. also offers a free, online (browser-based) file viewer that claims compatibility with over 50 file formats: https://viewer.autodesk.com/ accessed 18 June 2018<br><br>Robert McNeel & Associates currently offer downloadable evaluation versions of widely-used Rhinoceros 3D modelling software (versions 5 and 6) on their website, which continue to function for viewing and editing files after the end of the 90-day trial period (sans plugins and save/export functionality) thus enabling long-term access to the |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| carrying media (e.g. floppy disks) and ingest into the central archival repository. While commercially pressed CD-ROM disks may last indefinitely if cared for, consider making a backup copy and ingest into the central repository.<br><br>Share details of proprietary software holdings with other collecting institutions, and cooperate with other institutions to establish a network of legacy software users. | | proprietary .3dm file format. Rhinoceros is also able to import and parse (interpret) a wide variety of other proprietary file formats. https://www.rhino3d.com/download accessed 18 June 2018 |
| 4.3 Collect office style manuals, file naming and drawing conventions, document management and archival retention and disposal schedules, along with other relevant practice-wide documentation. | Style manuals, file naming and drawing conventions may be necessary for understanding the context and content of architectural records. | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| 4.4 Collect functioning legacy hardware, especially physical carrying media drives (e.g. tape drives, floppy and zip disk drives etc.).<br><br>Share details of functional legacy hardware holdings with other collecting institutions, and cooperate with other collecting institutions to establish a network of legacy hardware users.<br><br>Consider donating unwanted functioning legacy hardware to a collecting institution with a legacy hardware focus. | Legacy hardware is rapidly disappearing as organisations discard what they perceive to be obsolete technology.<br><br>Functioning mechanical hardware with moving parts and/or rubber belts that may degrade over time, e.g. tape drives, are especially rare. Such hardware may be valuable for other organisations looking to recover files stored on legacy media.<br><br>It is unlikely that every institution will be able to maintain all of the legacy hardware that they might need. | The Computer Archaeology Laboratory at Flinders University is an Adelaide-based organisation that collects and preserves functioning legacy hardware and actively and preserves legacy software. The Laboratory accepts donations: http://csem.flinders.edu.au/thegoodstuff/comparch/about.php accessed 18 June 2018<br><br>The Monash Museum of Computing History is based in Melbourne: https://www.monash.edu/it/about-us/museum-of-computing-history/about-the-museum accessed 18 June 2018<br><br>The Australian Computer Museum Society Inc. is based in Sydney: https://www.acms.org.au/ accessed 18 June 2018 |
| *5. Access* | | |
| 5.1 When access to CAD/BIM files through native software is not possible, open format derivatives of milestone CAD/BIM files should be created and ingested into the central archival repository alongside native format files.<br><br>5.2 CAD/BIM models should be converted (reformatted), where | Generating open format derivative versions of native CAD/BIM files, will ensure that at least some of the information (e.g. the basic geometry) will remain accessible into the future, even if access to proprietary software is unavailable.<br><br>Given the large number of CAD/BIM files that may be generated over the course of a project, generating open source derivatives of all CAD/BIM files may not be practical. However, if possible, converting milestone project files to open formats seems a reasonable compromise, since such files may be expected to meet the needs of most users (MacKenzie Smith 2009 p. 12). | The *NATSPEC National BIM Guide* provides guidelines on the types of files that should comprise milestone digital project deliverables (Construction and Information Systems Limited 2016): https://bim.natspec.org/documents/natspec-national-bim-guide accessed 14 June 2018<br><br>Alex Ball's DPC Technology Watch Report, *Preserving Computer-Aided Design (CAD)* provides details of open-source formats and outlines best-practice procedures for generating open-source derivatives of 3D CAD models. Ball also discusses special affordances of native 3D/BIM models that may be lost when converting into open source formats (Bath, UK: Digital Preservation Coalition 2013): http://dx.doi.org/10.7207/twr13-02 accessed 14 June 2018<br><br>MacKenzie Smith, *Final Report for the MIT FACADE Project: October 2006 – August 2009:* |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| necessary, into open-format derivatives on a case-by-case basis. Preservation action to facilitate access should be focused on a curated selection of CAD/BIM files that are expected to see the most use; other files should be preserved in perpetuity at the bit stream level. | Nevertheless, specialised research will likely benefit from access to a wider range of project records, therefore all native CAD/BIM files should be retained.<br><br>In the future, emulation will likely provide economical access to native CAD/BIM files that have been preserved at the bit stream level without intentional preservation actions targeted at providing long-term access to individual files. | https://www.cvaa.be/sites/default/files/projecten/bijlagen/bib_3896_facade_final.pdf accessed 18 June 2018 |
| 5.3 Consider a 'reformat on demand' approach to proprietary CAD/BIM files (analogous to 'digitisation on demand'). | | The National Library of Australia (NLA) offers a 'digitisation on demand' service for physical collection items: http://www.nla.gov.au/policy-and-planning/collection-digitisation-policy accessed 24 June 2018 |
| 5.4 Appropriate file formats for creating open derivatives may include for example:<br>• For word processing documents, PDF/A.<br>• For spreadsheets, CSV.<br>• For 2D CAD drawings and/or vector graphics, PDF/A and SVG.<br>• For bitmap images, TIFF.<br>• For 3D CAD formats, STEP, IGES and OBJ as appropriate.<br>• For BIM data, Industry Foundation Classes (IFC) for BIM models and PDF/A for 2D drawings. | | The *Sustainability of Digital Formats: Planning for Library of Congress Collection* website provides sustainability information on various digital file formats: http://www.loc.gov/preservation/digital/formats/fdd/descriptions.shtml accessed 24 June 2018<br><br>See also Ball, *Preserving Computer-Aided Design (CAD)* and the *NATSPEC National BIM Guide*. |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| *6. Metadata* | | |
| 6.1 Collect as much contextual metadata as possible for each project or collection, as appropriate, at the point of donation.<br><br>Contextual metadata may include, for example, name of architect, name of project, client, contractors, consultants, location, address, GPS coordinates, year designed, year completed, donated by, date donated, staff member received by, access conditions, etc.<br><br>Consider supplying donors with a pro forma to enable them to record information. | Contextual metadata is most easily captured at the point of donation, rather than at a later date, when donors may no longer be available to provide reliable information. | Dublin Core Metadata Initiative Metadata Terms: http://dublincore.org/documents/dcmi-terms/ accessed 27 June 2018<br><br>The *Categories for the Description of Works of Art* (CDWA) metadata schema developed by the J. Paul Getty Trust, 'provides options for data that closely resembles the "job" or project and individual document level found in design collections' (Waverly & Lowell 2006 p. 137): http://www.getty.edu/research/publications/electronic_publications/cdwa/ accessed 27 June 2018 |
| 6.2 File metadata (e.g. date created, date modified, date accessed, author etc.) can and should be generated automatically. | Recording file metadata manually is laborious, time-consuming, and prone to human error. | Digital metadata extraction tools are available for download from the Digital Curation Centre (DCC) website: http://www.dcc.ac.uk/resources/external/category/metadata-extraction accessed 27 June 2018 |
| 6.3 As a minimum, consider archiving within each project or collection folder:<br><br>a) A text document containing an outline of folder contents and documenting contextual metadata.<br><br>b) A spreadsheet containing a list of files and associated file metadata. | In the absence of a catalogue, database, or other collection management tool, including plain text contents documents in each project or collection folder, as appropriate, will help collection managers and/or users understand the structure and content of a digital collection. | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| *7. Retention and disposal (deletion)* | | |
| 7.1 Avoid weeding duplicate files unless absolutely necessary (e.g. duplicate copies of entire directory structures, or large files (e.g. 1 GB or larger).<br><br>If necessary to weed duplicate files, proceed with caution. | Weeding or culling duplicate copies of files may disrupt links and interdependencies between files.<br><br>Archivists should weight the costs of weeding duplicate files – in terms of time – and the potential for disrupting hidden links between files with the potential savings of file space that may be afforded by culling duplicates. Typically, the small savings of space to be achieved do not justify the time required (Greene and Meissner, 2005), except perhaps in the cases where entire directory structures are duplicated. | Mark A. Greene and Dennis Meissner, "More Product, Less Process: revamping traditional archival processing" in *The American Archivist,* Vol. 68 (Fall/Winter 2005) addresses processing backlogs in traditional (paper-based) archival processing, and offers a critique of fine-grained (i.e. item-level) preservation and description activities. http://www.archivists.org/prof-education/pre-readings/IMPLP/AA68.2.MeissnerGreene.pdf accessed 26 June 2018 |
| 7.2 Exercise caution when disposing of digital records.<br><br>Err on the side of retaining more rather than less.<br><br>"If in doubt, don't throw it out!" | Given that 'a decision not to preserve is usually a final one for digital materials,' the UNESCO *Guidelines for the Preservation of Digital Heritage* recommend a cautious approach to the disposal of digital records (p. 72).<br><br>The potential historical significance of architectural records may not be immediately obvious, and the future research value of digital architectural records (3D CAD and BIM models in particular) is difficult to predict.<br><br>All categories of records are likely to be of interest to future historians, including but not limited to, correspondence (emails, faxes), sketches and design development drawings and models, design intent and as-built CAD models and documentation, photographs, physical models, renderings, animations, presentations to clients (i.e. PowerPoint), presentation drawings, specifications, reports and contracts. | National Library of Australia, *Guidelines for the Preservation of Digital Heritage,* UNESCO (2003) http://www.unesco.org/new/en/communication-and-information/resources/publications-and-communication-materials/publications/full-list/guidelines-for-the-preservation-of-digital-heritage/ accessed 14 June 2016 |

## REFERENCES


Ball, Alex, *Preserving Computer-Aided Design (CAD),* DPC Technology Watch Report, Bath, UK: Digital Preservation Coalition (2013), http://dx.doi.org/10.7207/twr13-02 accessed 14 June 2018

Barrera-Gomez, Julianna and Ricky Erway, *Walk This Way: detailed steps for transferring born-digital content from media you can read in-house*, Dublin, Ohio: OCLC Online Computer Library Centre, Inc. (2013), http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf accessed 14 June 2018

Collins, Julie, Susan Collins and Christine Garnaut, "Behind the Image: assessing architectural drawings as cultural records" in *Archives and Manuscripts,* Vol. 35 No. 2 (November 2007) pp. 87-107

Digital Curation Centre (DCC), *Metadata Extraction,* http://www.dcc.ac.uk/resources/external/category/metadata-extraction accessed 27 June 2018

Digital Preservation Coalition, *Digital Preservation Handbook (2nd Edition)* (2015), https://www.dpconline.org/handbook accessed 27 June 2018

Dollar, Charles and Lori Ashley, *Digital Preservation Capability Maturity Model (DPCMM): Background and Performance Metrics Version 2.7* (July 2015), https://www.statearchivists.org/resource-center/resource-library/digital-preservation-capability-maturity-model-dpcmm/ accessed 14 June 2018

Dublin Core Metadata Initiative, *Metadata Terms:* http://dublincore.org/documents/dcmi-terms/ accessed 27 June 2018

Farneth, David, *Selecting, acquiring, managing, and preserving born digital collections… in 5 minutes!* https://www.ifla.org/files/assets/rare-books-and-manuscripts/Lyon-2014-presentations/wlic_2014_lyon_session_84_farneth.pdf accessed 27 June 2018

Greene, Mark A. and Dennis Meissner, "More Product, Less Process: revamping traditional archival processing" in *The American Archivist,* Vol. 68 (Fall/Winter 2005) http://www.archivists.org/prof-education/pre-readings/IMPLP/AA68.2.MeissnerGreene.pdf accessed 26 June 2018

J. Paul Getty Trust *, Categories for the Description of Works of Art* (CDWA), http://www.getty.edu/research/publications/electronic_publications/cdwa/ accessed 27 June 2018

J. Paul Getty Trust, *Categories for the Description of Works of Art,* http://www.getty.edu/research/publications/electronic_publications/cdwa/ accessed 27 June 2018

Lavoie, Brian, *The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition), UK: Digital Preservation* Coalition (2014), http://dx.doi.org/10.7207/twr14-02 accessed 24 June 2018

Leventhal, Aliza, *Designing the Future Landscape: Digital Architecture, Design & Engineering Assets,* A report on the Architecture, Design and Engineering Summit organized by the Library of Congress, the National Gallery of Art and the Architect of the Capitol on November 16 & 17, 2017 at the Library of Congress: http://digitalpreservation.gov/meetings/ade/ade2017.html accessed 20 June 2018

Library of Congress, *Sustainability of Digital Formats: Planning for Library of Congress Collection* website: http://www.loc.gov/preservation/digital/formats/fdd/descriptions.shtml accessed 24 June 2018

Lowell, Waverly and Tawny Ryan Nelb, *Architectural Records: managing design construction records,* Chicago, Ill: Society of American Archivists (2006)

Lynn, Greg (ed.) *Archaeology of the Digital,* Montreal: Canadian Centre for Architecture and Sternberg Press (2013)

National Archives, The (UK) *Preserving Digital Records: Guidance,* http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/guidance/ accessed 27 June 2018

National Library of Australia, "Mediapedia" https://mediapedia.nla.gov.au/home.php accessed 14 June 2018
National Library of Australia, *Guidelines for the Preservation of Digital Heritage,* UNESCO (2003), http://www.unesco.org/new/en/communication-and-information/resources/publications-and-communication-materials/publications/full-list/guidelines-for-the-preservation-of-digital-heritage/ accessed 14 June 2018

NATSPEC, *NATSPEC National BIM Guide*, Construction and Information Systems Limited (2016), https://bim.natspec.org/documents/natspec-national-bim-guide accessed 14 June 2018

Peyceré, David, "The Architectural Practices as First Curators of their Archives" NAi, Rotterdam. Hybrid Architectural Archives Conference (11-12 June 2009), conference.nai.nl/mmbase/attachments/525844/1.1_David_Peyceré.pdf accessed 14 June 2018

Phillips, Megan Jefferson Bailey, Andrea Goethals and Trevor Owens, *The NDSA Levels of Digital Preservation: An explanation and uses,* Library of Congress, National Digital Stewardship Alliance (2013), www.digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf accessed 16 May 2018

RLG/OCLC Working Group on Digital Archive Attributes, *Trustworthy Digital Repositories: Attributes and Responsibilities,* Mountain View, California: RLG (May 2002), www.digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf accessed 14 June 2018

Rosenthal, David S. H. Thomas Robertson, Tom Lipkisi, Vicky Reich and Seth Morabito, "Requirements for Digital Preservation Systems: A Bottom-Up Approach" in *D-Lib Magazine,* Vol. 11 No. 11 (November 2005), www.dlib.org/dlib/november05/rosenthal/11rosenthal.html accessed 14 June 2018

Rosenthal, David S. H., *Emulation & Virtualization as Preservation Strategies,* Andrew W. Mellon Foundation (2015), https://mellon.org/Rosenthal-Emulation-2015/ accessed 20 July 2018

Smith, MacKenzie, *Final Report for the MIT FACADE Project: October 2006 – August 2009:* https://www.cvaa.be/sites/default/files/projecten/bijlagen/bib_3896_facade_final.pdf accessed 18 June 2018

Svrcek, Ivan, *Accelerated Life Cycle Comparison of Millenniata Archival DVD,* China Lake, California: Life Cycle and Environmental Engineering Branch Naval Air Warfare Center Weapons Division (2013), https://archive.org/details/ChinaLakeFullReport accessed 14 June 2018

Walsh, Tim 'Catching up with the present: Archiving born-digital records of architecture and design' presentation at *Born digital: a symposium exploring Born digital: a symposium exploring digital architectural and built environment records,* Adelaide, 18 – 19 April 2016: http://www.unisa.edu.au/born-digital accessed 18 June 2018

Zaluendo, Maria Ines, 'Paradigm Shift: curatorial views on collecting and archiving architectural drawings in an evolving born-digital landscape' paper presented at the Society of American Archivists Conference, Washington D.C. (August 2014) p. 3, https://dash.harvard.edu/bitstream/handle/1/13442962/IZ_Born-Digital%20Architectural%20Drawings_FINAL_2014_A1b.pdf accessed 17 May 2018

# A draft national framework for archiving digital architectural records in Australian architectural practice

# v.1.0 (June 2018)

## Introduction

The emergence of digital technology has had a significant impact on the way in which buildings are designed and constructed. From being regarded initially as a tool to aid design, the computer is now commonly considered to be integral to the design process. The digital environment in which an architectural project is developed involves computer hardware and software in the creation of digital files. The records of the process of designing a building cover a broad spectrum. They include models that explore its potential shape and form, sketches, plans, elevations, sections, renderings and other documents like photographs, emails, letters, faxes, specifications and contracts. Increasingly, these records are produced in digital environments and only exist as digital files.

Currently in Australia a framework does not exist to guide collecting, archiving and preserving these “born digital” architectural records. As a consequence, their future preservation as historical and cultural records and their accessibility to practitioners and researchers is in jeopardy. Architects in practice and archivists in collecting institutions face similar challenges when endeavouring to preserve digital architectural records. The development of this framework was informed by these challenges which are captured in the following four questions:

1. What rationales exist for archiving digital architectural records?
2. Which records should be archived?
3. How should records be archived?
4. How can digital continuity be achieved in rapidly evolving and changing electronic environments?

The framework focuses on digital preservation, and includes:

- A glossary of technical terminology.
- A statement of preservation intent for digital architectural records, describing the different categories of digital architectural records and the significant properties of each category that should be preserved.
- An outline of digital preservation fundamentals, written for a non-specialist audience. The outline stipulates four requirements for successful digital preservation, and concludes with practical advice for individuals or organisations with no or limited IT expertise.
- A set of guidelines for digital preservation, covering several areas of concern: bit-level preservation; physical carrying media; file formats, software and hardware; naming and filing procedures; and retention, disposal and donation schedules.
- A table that expands upon the guidelines, providing a rationale for each guideline and relevant resources and further reading that may be useful for the practical implementation of the guidelines.
- A list of references.

A companion document, *A draft national framework for archiving digital architectural records in Australian collecting institutions – v.1.0 (June 2018),* is targeted at professional architects. Archivists may benefit from reading this document; likewise, architects may benefit from a deeper understanding of the rationales that underlie the decision-making process of archivists when selecting architectural records for preservation in perpetuity. These two documents complement and reinforce each other.

The preparation of this framework was funded by NATSPEC.

## Glossary

| | |
|---|---|
| BIM | Building Information Model. A 3D *CAD* model that contains embedded information on materials, cost estimation and scheduling. |
| Bit rot | slow and inevitable degradation of a bit stream over time. |
| Bit stream | a sequence of *bits* that stores digital information. |
| Bit | the basic unit of digital information. A bit can have one of only two possible values (on/off), most commonly represented as 0 or 1. |
| CAD | Computer Aided Design. Software used by designers to create precision 2D technical drawings and 3D models. |
| Catastrophic failure | a sudden and total failure from which recovery is impossible. |
| CD | Compact Disk, an *optical physical carrying media* |
| Checksum | a short sequence of characters that acts as a fingerprint of a digital file. A checksum is generated by passing the *bit stream* of a digital file through an algorithm. A checksum may be used to *validate* the *fixity* of a *bit stream*. |
| Cloud computing | the delivery of hosted services over the internet. |
| Corruption | an unintended change in the *bit stream* of a digital file. |
| Dark storage | digital information stored on physical carrying media that is not connected to a computer network. |
| Directory | the structure of a computer filing system; a hierarchy of files and folders. |
| Disk imaging | the process of creating an identical copy of the data stored on a unit of physical carrying media, including empty space and deleted files. |
| DVD | Digital Versatile Disk, an *optical physical carrying media* |
| Emulation | using one computer system to imitate the functions of another. |
| Fixity | (of a bit stream) the state of being unchanging or permanent. |
| Flash memory | digital information stored on a computer chip that can be electrically erased and reprogrammed. |
| Hard disk drive | a *physical carrying media* technology in which large quantities of digital data is stored on a rapidly rotating magnetic platter. |
| Hardware lock | a physical device that renders *software* inoperable when not physically connected to the computer. |

| | |
|---|---|
| Hardware | the physical components of a computer, in contrast to *software*, which are programs or instructions that can be run by hardware. |
| IGES | Initial Graphics Exchange Specification. An open file format for storing and/or exchanging CAD data. |
| Ingest | the process of absorbing content into an archive. |
| Installation files | files that allow the installation of a piece of *software* on a computer system. |
| IT | Information Technology |
| JPEG | Joint Photographic Experts Group. A file format for storing *raster graphics* images. |
| Legacy | (of a technology) no longer in current or regular use. |
| Lossless | (of copying or compression) without loss of information. |
| Metadata | data about data; data attached to files stored on a computer system, for e.g. author, date modified, date created. |
| Native format | the format that a piece of software uses by default. |
| OBJ | an *open file format* for storing 3D models. |
| Open format | a file format that is not protected by proprietary patents or other intellectual property restrictions. |
| Optical media | optical *physical carrying media*, in which digital data is stored as markings on disk. The disk is read by a laser while spinning at a high speed. |
| PDF-A | Portable Document Format Archival. An *open file format* for the long-term storage of document content, including text and images. |
| Physical carrying media | the physical media that are used to store information, for example, hard disk drives, magnetic tape or disks, flash media, optical disk (for example CD/DVD recordable) etc. |
| Proprietary format | a file format that is protected by patents or other intellectual property restrictions. |
| Raster graphics | Digital images made up of a rectangular grid of pixels. Compare *vector graphics.* |
| Server | an array of hard disk or solid state drives accessed via a local computer network. |
| Software | programs or instructions that tell a computer how to work, in contrast to physical *hardware* that performs work. |
| Solid state drive | a physical carrying media technology in which large quantities of digital data is stored using *flash memory*. |
| STEP | Standard for the Exchange of Product Model Data. An *open file format* for storing *CAD* data. |

| | |
|---|---|
| SVG | Scalable Vector Graphics. An *open file format* for storing *vector graphics* images. |
| TIFF | Tagged Image File Format. An *open file format* for storing *raster graphics* images. |
| Validate | to check the *fixity* of digital files using *checksums.* |
| Vector graphics | Digital images composed of paths or vectors. Compare *raster graphics.* |

## Statement of preservation intent for born digital architectural records

Digital architectural archives may include any number of file types, including but not limited to: 2D documents, including proprietary word processing, spreadsheet, desktop publishing and presentation files; 2D images, including proprietary format raster and vector image editing files, and open format images; proprietary native format 2D and 3D CAD and BIM models and open format CAD and BIM models (including PDF drawings), email archives, video files, audio files and other proprietary software files. Digital architectural archives will also likely contain proprietary and open source software.

This statement of preservation intent addresses physical carrying media, directory structures and interrelationships, original file metadata, and three common categories of digital objects that are likely to be found in digital architectural collections, which are also most likely to be problematic for maintaining long-term access: 2D images and documents, native file format CAD and BIM models, and CAD and BIM software.

### Physical carrying media, file directories, and metadata

The bulk of born-digital architectural records, not unlike their physical counterparts, are typically organised according to the architectural projects for which they were created. Born digital architectural records are often organised into hierarchical groups of files and folders or *project directories.* Project directories may have evolved organically as the project unfolded, or may be the result of an order imposed on a group of files and folders later.

Sometimes project files may be split across multiple carrying media. Sometimes files may be split across multiple carrying media in a proprietary archive format (i.e. a ZIP format), requiring access to proprietary software to access files in the archive.

Occasionally *working directories* are archived – essentially a copy of the designer's desktop, containing whatever files they were working on at the time.

Born-digital architectural records that may fall outside project directories, but which may be necessary for understanding the context of records within project directories, may include for example office style manuals, file naming and drawing conventions, document management guidelines and archival retention and disposal schedules and email and/or correspondence archives.

The relationship between files, whether in a project or working directory structure, or when split across carrying media, should be retained. Altering the original directory structure of a project carries with it the likelihood of disrupting links and other interdependencies between files. For this reason, weeding or culling duplicate files may be problematic. Furthermore, the original directory structure – however disordered – may reveal important contextual information about the files themselves, or the working practices of a firm at a particular point in time, which may be of interest to future researchers.

File metadata (e.g. dates created, modified etc.) may be essential to identifying, for example, the most recent version of a file or establishing the chronological sequence of work. It is essential that original file metadata is at least recorded, but is best left unaltered.

Best practice for preserving directory structure and file metadata is to create forensic disk images of physical carrying media, and to preserve disk images in perpetuity as an enduring record of files in their original state, as near as possible to the point of donation. File metadata may also be extracted from files automatically into a spreadsheet using special software.

Generally speaking, the files and directory structure are important, not the physical carrier. However, the physical carrier may be labelled with metadata and should be documented. In some cases, the physical carrier may be significant in itself as an example of a legacy technology, perhaps illustrating the working practices within a particular firm.

*Summary:*

- Relationships between files in a project directory, in a working directory or across carrying media – however disordered – should be maintained.
- Preserve original file metadata.
- Avoid weeding or culling duplicate files unless there are sound reasons for doing so.
- Create disk images of physical carrying media and retain them in perpetuity.
- Retain physical carrying media where possible. At minimum, document metadata on physical carrying media before disposal.

**Two dimensional (2D) documents and images**

2D documents may include files in proprietary formats including, for example, Microsoft Word, Excel and PowerPoint files. 2D documents may also include open formats like PDF, including PDF derivatives of drawings created in CAD software.

2D images may include open format raster (e.g. TIFF, JPEG, PNG) and vector images (e.g. SVG), as well as proprietary editing formats created in, for example, Adobe Photoshop, Adobe Illustrator or Adobe InDesign.

In general, the ability to edit 2D images and documents in proprietary formats is not required, and open format derivatives may be created to facilitate long-term preservation and access. However, it is possible that CAD/BIM files may have hidden interactive links to some 2D documents or images, which may be lost if original proprietary format files are discarded – for example, images may be linked to CAD/BIM files as textures applied to 3D geometry. Therefore, reformatting of some image files into different formats without retaining the originals in their original context may be problematic; for example, changing the file extension may cause a link to become severed.

Preservation action to facilitate access could be focused on a curated selection of 2D documents and images that are expected to see the most use; all other files should be preserved in perpetuity at the bit stream level, even if access to proprietary software is temporarily lost.

*Summary:*

- Create open format derivatives of proprietary formats as required, focusing on a small selection of files that will satisfy the needs of most archive users.
- Preserve all proprietary files in their original context in perpetuity at the bit stream level.

**Native file format CAD and BIM models**

Viewing CAD/BIM models in their original software environments is highly desirable, since many aspects of a CAD/BIM model's functionality can only be experienced by opening the model in the original software environment. This is especially true of 3D CAD and BIM models, which have no analogue equivalents, but may also be true of 2D models or drawings. 3D CAD and BIM models contain information that is embedded nowhere else – not even in the completed building.

Nevertheless, viewing CAD/BIM models in their original software environments may not be possible since CAD/BIM models in native formats are usually proprietary and rely on access to proprietary software, which may be dependent on legacy hardware and/or licencing.

However, the preservation intent for CAD/BIM models is that native software files are preserved in perpetuity, even though access to the content of files through proprietary software may be temporarily lost.

It is expected that in the future, access to some legacy file formats may be regained through software emulation. It is also possible that certain proprietary file formats may be released into the public domain when their patents expire (for example the MP3 format, formerly protected by patents, and now an open format). Additionally, other collecting institutions or even individual archive users may possess copies of the proprietary software needed to access the content of files, even if the collecting institution preserving the files does not.

Since access to the content of native CAD/BIM models is precarious, CAD/BIM models should be converted (reformatted) into open-format derivatives on a case-by-case basis. While each architectural project may contain a large number of CAD/BIM files, access to a small number of files from each project (for example the final design intent model, or the final as-built model) is expected to serve the needs of most users. Preservation action to facilitate access could be focused on a curated selection of CAD/BIM files that are expected to see the most use; other files should be preserved in perpetuity at the bit stream level. All native proprietary files should be maintained in perpetuity, even if access to native software is temporarily lost (loss of access to proprietary software and file formats should always be considered temporary).

Where open source derivatives already exist side-by-side with native CAD files in a project directory, both should be retained in perpetuity.

*Summary:*

- Access to proprietary CAD/BIM files in their native software environments is desirable but not always possible.
- Preserve all native CAD files in perpetuity at the bit stream level.
- Create open format derivatives for a small selection of native CAD files that will satisfy the needs of most archive users.
- Where open source derivatives already exist side-by-side with native CAD files and a project archive, both formats should be retained in perpetuity.

**CAD and BIM software**

Because collecting digital architectural records means collecting files in proprietary formats which are in many cases dependent on proprietary software, preserving digital architectural records means preserving the proprietary software that was used to create them. Software includes for example installation files on physical media, manuals, boxes and documentation, certificates of authenticity, licence keys and hardware locks.

All CAD/BIM software should be preserved in perpetuity, even if not supported by present hardware. As discussed above, it is expected that in the future access to some legacy software may become possible through emulation.

Physical documentation and manuals may be retained in hardcopy form, or digitised, or both. The content of documentation and/or manuals is important, not the physical format.

Certificates of authenticity should be retained in physical form.

Hardware locks are essential to operating some software and must be retained in physical form.

When installation files exist on volatile carrying media (e.g. floppy disks) they may need to be ingested into the central archival repository to receive bit stream-level preservation. Files carried on commercially pressed CD-ROM disks may be ingested into the central archival repository, however such media will last indefinitely if cared for, and so may be retained as well or instead.

*Summary:*

- Software includes installation files on physical media, manuals, boxes and documentation, certificates of authenticity, licence keys and hardware locks.
- Preserve all software installation files in perpetuity at the bit stream level.
- Retain a copy of all documentation, manuals and software boxes – either in physical or digitised form, or both.
- Retain certificates of authenticity in physical form.
- Retain hardware locks in physical form.
- Ingest all installation files on volatile physical carrying media into the central archival repository for bit stream level preservation in perpetuity.
- Retain commercially pressed CD-ROM disks, or ingest file contents into the central archival repository for bit stream level preservation in perpetuity, or both.

## Digital preservation fundamentals

All digital files are composed of *bits* – the smallest possible units of digital information. Bits are usually represented as 0s and 1s. A sequence of bits is called a *bit stream*. A bit stream spells out information in a *binary code* that is understandable by computers.

Digital files may be copied an infinite number of times without degradation – they may be said to be *lossless.* Analogue media, on the other hand, will degrade with each *generation* of copying (for example a photocopy will be of a lower quality than the original; a photocopy of a photocopy will be of a lower quality still). Note that some digital file formats have lossy *encoding* of their *content* (e.g. JPEG image) where digital information is discarded at the point of encoding (i.e. the act of "saving as" a JPEG discards information forever). Nevertheless, the bit stream of a JPEG image, or any other file, can be copied an infinite number of times without *further* degradation.

The media on which a digital bit stream is stored is known as *physical carrying media* – 'physical' because carrying media exists physically, in the real world; 'carrying' because it stores or 'carries' digital *bits.* Examples of physical carrying media include, for example, hard disk drives, solid state drives, magnetic media (e.g. floppy disks, magnetic tape) optical media (e.g. CD/DVD/Blu-ray recordable disks) and flash media (e.g. digital camera cards, USB thumb drives).

A *server* is usually an array of hard disk or solid state drives, accessed via a local computer network. *Cloud storage* is usually a networked array of hard disks or solid state drives, operated by a third party and accessed via the internet.

Depending on the physical carrying media, bits may be stored magnetically, or as electrical charges, or as physical indentations on the surface of a disk, or holes in a punched card, among other methods.

All currently available physical carrying media have comparatively short reliable lifespans compared with hardcopy (paper-based) records, typically within the range of 2-5 years.

All currently available physical carrying media are susceptible to unpredictable *catastrophic failure*, when access to all stored data may be suddenly lost, without warning.

All currently available physical carrying media is susceptible to slow degradation, where individual or multiple bits may spontaneously change over time – for example a 0 may change into a 1, or a 1 may change into a 0, or a bit may be omitted, or an extra bit may be added, or an entire string of bits may change.

Once a bit in a digital file has changed, the file is said to be *corrupted.* Once a digital file has corrupted, it is practically impossible to repair.

While corrupted files cannot be repaired, the lossless nature of digital information makes it easy to prepare identical copies of digital files.

Identical copies should be kept in different *physical locations*, that is, not together on the same physical carrying media, which may fail catastrophically.

If a file becomes corrupted, it is easy to replace it with an identical copy.

The integrity or *fixity* of digital files must be monitored so that corruption may be detected and corrupted files replaced with identical copies.

A *checksum* is a short sequence of characters that acts as a fingerprint of a digital file. A checksum is *generated* by passing the bit stream of the digital file through an *algorithm*. If the bit stream becomes corrupted, *even by a single bit,* a different checksum will be generated when the bit stream is passed through the checksum algorithm.

Periodically comparing a digital file to its corresponding checksum will reveal whether or not a digital file has become corrupted.

Checksums can be generated automatically for large numbers of digital files.

Large numbers of digital files can be automatically checked against checksums generated earlier.

**Therefore, preservation of digital bit streams requires:**

1. Maintaining at least *two identical copies* of all digital files in *different physical locations*.

2. Generating and keeping *checksums for all copies* of all digital files.

3. *Checking* the fixity of all copies of digital files *regularly* against checksums.

4. *Replacing corrupted files* when detected with identical copies.

*If these four essential bit stream-level digital preservation actions are not observed, all other digital preservation actions are futile.*

Checksums can be easily generated and checked for the contents of entire *directories* (a hierarchy of files in folders) with freely available, user-friendly utilities such as TeraCopy: http://www.codesector.com/teracopy (accessed 23 June 2018).

If the above preservation actions cannot be applied, it is suggested that individuals with limited information technology expertise do the following:

1. Burn at least two copies of all files (preferably with checksums) to M-DISC DVD-R recordable media.

2. Burn at least two copies of all files (preferably with checksums) to regular DVD-R recordable media, using at least two different disk brands (e.g. TDK and Verbatim).

3. Label and date all disks on the central (usually transparent) hub.

4. Store one copy on each media in a separate place (i.e. one copy on M-DISC and one copy on regular DVD-R in one place, the other copies in another place).

5. Complement each disk with an identical copy of that disk every five years. Do not discard the original disks.

6. Label and date complementary copies.

*Note that these steps will not guarantee the preservation of digital data, but they will increase the likelihood that digital data will survive until it falls into the custody a preservation repository.*

See also David Farneth, *Selecting, acquiring, managing, and preserving born digital collections… in 5 minutes!* https://www.ifla.org/files/assets/rare-books-and-manuscripts/Lyon-2014-presentations/wlic_2014_lyon_session_84_farneth.pdf accessed 27 June 2018.

**Key Guidelines**

*1. Bit-level preservation*

1.1. A designated individual within an architectural practice should have a basic understanding of bit-level digital preservation principles. The duties of archive manager should be allocated to this individual.

1.2 Data for completed (archived) project files should be ingested into a central archival repository.
This repository should be separate from the working drive of the practice and could take several forms, e.g.:

- A networked server (for large practices)
- A dedicated hard disk or solid state drive (for small practices)
- An external storage provider

1.3 At minimum two complete copies of the central archival repository should be maintained. One copy should be maintained offsite.
Best practice: maintain at least three complete copies of the central archival repository. One copy should be maintained in a geographically distant location.

1.4 An external storage provider may be used to maintain an offsite copy of the central archival repository.
Best practice: external storage providers should be ‘trustworthy’ (ISO 16363 / TDR certified or similar) and provide bit stream-level preservation of digital files.

1.5 All files in the central repository should have associated fixity information. All files in the repository should be checked against fixity information on a regular basis. Corrupted files should be replaced with copies prepared at the time of archiving.

1.6 Files in backup copies of the repository should also be checked against fixity on a regular basis. Corrupted files should be replaced from the main repository or another backup repository.

*2. Physical carrying media*

2.1 Data for completed (archived) files that only exists on disparate carrying media (e.g. floppy disks, optical disks, CD, DVD and Blu-ray recordable disks, old hard disk drives, flash media etc.) should be ingested into the central archival repository.

2.2 Completed (archived) project data ingested into the archive from disparate carrying media should be copied in such a way that the original directory structure and metadata are retained intact.

2.3 Exercise care when copying files from carrying media.
Minimum practice: validate files on carrying media against copies ingested into the repository.
Best practice: generate disk images of physical carrying media, and ingest disk images into the digital archive alongside copies of validated files.

2.4 Exercise caution when disposing of original physical carrying media.
Minimum practice: retain corrupted media for future forensic recovery.
Best practice: retain all physical carrying media.

*3. File formats, software and hardware*

3.1 At minimum, open format derivatives of milestone CAD/BIM files should be created and ingested into the central archival repository alongside native format files.
Best practice: generate at least two different open format versions of all CAD/BIM files and archive alongside native format files.

3.2 Appropriate file formats for creating open derivatives of BIM data include the Industry Foundation Classes (IFC) format for BIM models and PDF/A for 2D drawings.

3.3 For non-BIM workflows, appropriate open file formats include for example, PDF/A and SVG for 2D CAD drawings and/or vector graphics, TIFF for bitmap images, and STEP, IGES and OBJ for 3D CAD formats.

3.4 For CAD/BIM files that will be retained in the central archival repository, always retain native CAD/BIM files along with any open format derivatives of the native files.

3.5 Archive a copy of office style manuals, file naming and drawing conventions, document management and archival retention and disposal schedules, along with other relevant practice-wide documentation alongside each project in the central archival repository.

3.6 Archive as much software as possible, including installation media, manuals and documentation, licence keys and hardware locks.
Ingest copies of installation files into the central archival repository. Where necessary, transfer installation files from legacy carrying media (e.g. floppy disks) and ingest into the central archival repository.
While commercially pressed CD-ROM disks will last indefinitely if cared for, consider making a backup copy and ingest into the central repository.

3.7 Retain functioning legacy hardware, especially physical carrying media drives (e.g. tape drives, floppy and zip disk drives etc.) for as long as they are required for accessing and transferring files stored on physical carrying media.
When no longer required, offer functioning legacy hardware as a donation to a collecting institution.

*4. Naming and filing procedures*

4.1 At minimum, file names should include a short specific plain English title, a date (possibly as a six or eight-digit code e.g. 27062018) and a variant or version number (e.g. v1_1 means variant 1 version 1).
Avoid 'illegal' characters and symbols in file names.
Use folder names to assign context to groups of files.

4.2 Consider implementing a quality management system guided by ISO 9000; certification need not necessarily be the final goal.

*5. Retention, disposal and donation schedules*

5.1 Architectural practices should develop a policy documenting the types of records (analogue and digital) that will be archived, as well as retention and disposal schedules for the records in their care.
This policy should be revisited and revised on a scheduled basis (e.g. once every five years).
Record and retain a list of what has been destroyed (paper records) or deleted (digital records)

5.2 While developing retention and disposal schedules, architectural practices should address the question of donating records to collecting institutions (libraries, archives or museums).
Intentions to donate records (which records will be donated, when and any conditions on public access) should be documented in schedules of retention and disposal.

5.3 While architects have traditionally donated records near the ends of their careers, donating digital architectural records should be donated earlier rather than later.

5.4 Exercise caution when disposing of records, especially digital records.
If practices intend to donate records to collecting institutions in the future, err on the side of retaining more rather than less.
"If in doubt, don't throw it out!"

5.5 As a minimum, practices should retain the following records for each project: 1) milestone project deliverables (outlined in more detail in 5.6), including as-built and design intent contract documentation and specifications; 2) a selection of design development records (including e.g. sketches and studies); 3) photographs, especially construction progress photographs.

5.6 Retain at least a minimum of records for unbuilt projects.

5.7 Set aside time at the end of a project for deliberately organising and archiving project records.

5.7 As outlined in the NATSPEC National BIM Guide, milestone digital project deliverables are to be submitted to clients and other stakeholders on CD/DVD recordable media with data clearly organised and software versions labelled.

Digital deliverables for BIM workflows are outlined in the NATSPEC National BIM Guide (pp. 19-20).

For non-BIM workflows, in lieu of BIM construction coordination and design intent models, digital deliverables should include final versions of all 2D and 3D CAD files in both native CAD formats and at least one open format.

Best practice: submit milestone digital project deliverables to clients and other stakeholders on M-DISC CD/DVD recordable media.

Milestone digital deliverables should also be ingested into the central archival repository in a labelled folder with data clearly organised and software versions labelled.

**Key Guidelines: Rationale, Resources and Further Reading**

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| *1. Bit-level preservation* | | |
| 1.1. A designated individual within an architectural practice should have a basic understanding of bit stream-level digital preservation principles. The duties of archive manager should be allocated to this individual. | Bit stream-level preservation is the foundation for all digital preservation activities; if the bit-stream fixity of digital objects cannot be assured, all other digital preservation actions are futile.<br><br>All commercially available digital carrying media are subject to both slow degradations over time ('bit rot') and the possibility of sudden, catastrophic failure.<br><br>Even if not directly responsible for bit-level preservation infrastructure, cultivating an understanding of bit-level preservation principles will enable architects to confidently discuss their digital preservation requirements with IT support personnel or external digital storage providers. | The NDSA *Levels of Digital Preservation* offers a set of plain-language, rubric-based guidelines for digital preservation: www.digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf accessed 14 June 2018<br><br>The *Digital Preservation Capability Maturity Model (DPCMM)* provides digital preservation performance metrics benchmarked against ISO 14721 conformance: https://www.statearchivists.org/resource-center/resource-library/digital-preservation-capability-maturity-model-dpcmm/ accessed 14 June 2018<br><br>David S. H. Rosenthal, Thomas Robertson, Tom Lipkisi, Vicky Reich and Seth Morabito, "Requirements for Digital Preservation Systems: A Bottom-Up Approach" in *D-Lib Magazine* (Vol. 11 No. 11, November 2005) provides an in-depth discussion of the challenges of digital preservation: www.dlib.org/dlib/november05/rosenthal/11rosenthal.html accessed 14 June 2018<br><br>Digital Preservation Coalition, *Digital Preservation Handbook (2nd Edition)* (2015) https://www.dpconline.org/handbook accessed 27 June 2018<br><br>The National Archives (UK) *Preserving Digital Records: Guidance* http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/guidance/ accessed 27 June 2018 |
| 1.2 Data for completed (archived) project files should be ingested into a central archival repository.<br><br>This repository should be separate from the working drive of the practice and could take several forms, e.g.:<br>• A networked server (for large practices)<br>• A dedicated hard disk or solid state drive (for small practices)<br>• An external storage provider | Centralising files in one location streamlines the process of checking the fixity of digital files (see below). Storing digital files across multiple physical carrying media (for example a series of CD recordable disks) will make fixity checking a tedious, manual process. | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| 1.3 At minimum two complete copies of the central archival repository should be maintained. One copy should be maintained offsite.<br><br>Best practice: maintain at least three complete copies of the central archival repository. One copy should be maintained in a geographically distant location. | Complete copies of the repository are maintained so that corrupted files may be replaced with identical copies.<br><br>A magnetic tape backup that is regenerated on a regular basis from the central repository could serve as one copy, but is still subject to degradation over time and should not be trusted as long-term 'dark storage'. | |
| 1.4 An external storage provider may be used to maintain an offsite copy of the central archival repository.<br><br>Best practice: external storage providers should be 'trustworthy' (ISO 16363 / TDR certified or similar) and provide bit stream-level preservation of digital files. | External storage providers may not necessarily provide secure, bit-level preservation of digital files.<br><br>Using an ISO 16363 certified provider increases the likelihood that the original bit stream of digital files submitted for long-term storage will be preserved. | ISO International Standard 16363 describes criteria for the audit and certification of trustworthy digital repositories. |
| 1.5 All files in the central repository should have associated fixity information. All files in the repository should be checked against fixity information on a regular basis. Corrupted files should be replaced with copies prepared at the time of archiving. | Checking against previously generated fixity information is the only efficient and economical way of monitoring the integrity of files. Note that fixity information can be generated automatically for large numbers of files. Likewise, large numbers of files can be checked against fixity information automatically. Once corruption has occurred, files can only be recovered by replacing them with backup copies. | For examples of software able to generate checksums and validate files see Julianna Barrera-Gomez and Ricky Erway, *Walk This Way: detailed* |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| 1.6 Files in backup copies of the repository should also be checked against fixity on a regular basis. Corrupted files should be replaced from the main repository or another backup repository. | Backup copies of the repository will degrade over time just as the main repository will degrade over time. | *steps for transferring born-digital content from media you can read in-house* (Dublin, Ohio: OCLC Online Computer Library Centre, Inc., 2013): http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf accessed 14 June 2018 |
| *2. Physical carrying media* | | |
| 2.1 Data for completed (archived) files that only exists on disparate carrying media (e.g. floppy disks, optical disks, CD, DVD and Blu-ray recordable disks, old hard disk drives, flash media etc.) should be ingested into the central archival repository. | Centralising files in one location streamlines the process of checking the fixity of digital files. Checking the fixity of files stored on disparate carrying media is laborious and time consuming. CD/DVD recordable media is especially vulnerable to damage when handled, resulting in possible data loss. | For practical advice on transferring files stored on disparate physical carrying media, including the use of write blockers and the generation of disk images, as well as links to software, see *Walk This Way: detailed steps for transferring born-digital content from media you can read in-house,* http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf accessed 14 June 2018<br><br>The Digital Preservation team at the National Library of Australia (NLA) has developed an online, searchable database of carrying media, called Mediapedia, designed to assist with the identification of various physical carrier formats. See: https://mediapedia.nla.gov.au/home.php accessed 14 June 2018<br><br>Tim Walsh, 'Catching up with the present: Archiving born-digital records of architecture and design' presentation at *Born digital: a symposium exploring Born digital: a symposium exploring digital architectural and built environment records,* Adelaide, 18 – 19 April 2016: http://www.unisa.edu.au/born-digital accessed 18 June 2018 |
| 2.2 Completed (archived) project data ingested into the archive from disparate carrying media should be copied in such a way that the original directory structure and metadata are retained intact. | Altering the original directory structure of a project carries with it the likelihood of disrupting links and other interdependencies between files. Furthermore, the original directory structure – however disordered – may reveal important contextual information about the files themselves or the working practices of a firm at a particular point in time, which may be of interest to future historians. Likewise, metadata (e.g. dates created, modified etc.) may be essential to identifying, for example, the most recent version of a file or establishing the chronological sequence of work. | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| 2.3 Exercise care when copying files from carrying media.<br><br>Minimum practice: validate files on carrying media against copies ingested into the repository.<br><br>Best practice: generate disk images of physical carrying media, and ingest disk images into the digital archive alongside copies of validated files. | Files may be incorrectly or incompletely transferred when relying on the built-in copy/paste functionality of a computer's operating system.<br><br>Disk imaging may allow data on corrupted physical media to be partially recovered. | |
| 2.4 Exercise caution when disposing of original physical carrying media.<br><br>Minimum practice: retain corrupted media for future forensic recovery.<br><br>Best practice: retain all physical carrying media. | Files may be incorrectly or incompletely transferred when relying on the built-in copy/paste functionality of a computer's operating system.<br><br>Improvements in technology may render physical media previously thought to be corrupted beyond repair accessible once again.<br><br>The physical carrier may be significant in itself as an example of a legacy technology, perhaps illustrating the working practices within a particular firm. | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| *3. File formats, software and hardware* | | |
| 3.1 At minimum, open format derivatives of milestone CAD/BIM files should be created and ingested into the central archival repository alongside native format files.<br><br>Best practice: generate at least two different open format versions of all CAD/BIM files and archive alongside native format files. | Generating open format derivative versions of native CAD/BIM files will ensure that at least some of the information (e.g. the basic geometry) will remain accessible into the future, even if access to proprietary software is lost.<br><br>Given the large number of CAD/BIM files that may be generated over the course of a project, generating open source derivatives of all CAD/BIM files may not be possible; however, milestone project files should as a minimum be converted into open source formats.<br><br>While open format derivatives are a useful insurance for 'future-proofing' architectural records, native CAD/BIM files are the "originals" and contain information that may be lost when converting them into open format derivatives, for example:<br>• Construction history that may be rolled back and forward like an interactive "undo/redo" feature. The history of modelling actions may be rolled back to an earlier point in the model's development, changes implemented, and the model "rolled forward" again, automatically updating later stages of construction with the results of the changes.<br>• Parametric modelling, in which 'aspects of the design are given a variable value instead of a fixed one, in order to make them easier to adjust and reuse in different contexts' with 'constraints [that] control how designs should be adjusted in the light of changed variables' (Ball p. 33). | The *NATSPEC National BIM Guide* provides guidelines on the types of files that should comprise milestone digital project deliverables (Construction and Information Systems Limited 2016): https://bim.natspec.org/documents/natspec-national-bim-guide accessed 14 June 2018<br><br>Alex Ball's DPC Technology Watch Report, *Preserving Computer-Aided Design (CAD)* provides details of open-source formats and outlines best-practice procedures for generating open-source derivatives of 3D CAD models. Ball also discusses special affordances of native 3D/BIM models that may be lost when converting into open source formats (Bath, UK: Digital Preservation Coalition 2013): http://dx.doi.org/10.7207/twr13-02 accessed 14 June 2018<br><br>MacKenzie Smith, *Final Report for the MIT FACADE Project: October 2006 – August 2009:* https://www.cvaa.be/sites/default/files/projecten/bijlagen/bib_3896_facade_final.pdf accessed 18 June 2018 |
| 3.2 Appropriate file formats for creating open derivatives of BIM data include the Industry Foundation Classes (IFC) format for BIM models and PDF/A for 2D drawings. | | |
| 3.3 For non-BIM workflows, appropriate open file formats include for e.g. PDF/A and SVG for 2D CAD drawings and/or vector graphics, TIFF for bitmap images, and STEP, IGES and OBJ for 3D CAD formats. | | |
| 3.4 For CAD/BIM files that will be retained in the central archival repository, always retain native CAD/BIM files along with any open format derivatives of the native files. | | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| | • Features, which are the parametric building blocks of a model within some solid modelling applications (for e.g. Solidworks and Autodesk Inventor) and in many cases have real-life manufacturing analogues – for example, “draft angle” (taper) added to a moulded part to allow it to be removed from a mould, a “fillet” (curved blended surface) to reduce stress where two surfaces meet, or a solid body hollowed out to create a “shell”.<br><br>Retaining native CAD/BIM files, even when access to proprietary software is lost, leaves the door open to future possibilities that may enable regained access, for example:<br>• Software emulation<br>• Archive users with active software licences<br>• Software vendors and/or patent holders releasing previously proprietary formats as open standards (as occurred with .mp3) | |
| 3.5 Archive a copy of office style manuals, file naming and drawing conventions, document management and archival retention and disposal schedules, along with other relevant practice-wide documentation alongside each project in the central archival repository. | Office style, drawing and file naming conventions may change over time. Archiving current conventions alongside project files at the conclusion of a project ensures that relevant office practices contemporary to the project are captured and archived. | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| 3.6 Archive as much software as possible, including installation media, manuals and documentation, licence keys and hardware locks.<br><br>Ingest copies of installation files into the central archival repository. Where necessary, transfer installation files from legacy carrying media (e.g. floppy disks) and ingest into the central archival repository.<br><br>While commercially pressed CD-ROM disks will last indefinitely if cared for, consider making a backup copy and ingest into the central repository. | Retaining native CAD/BIM files, even when access to proprietary software is lost, leaves the door open to future possibilities that may enable regained access, for example:<br>• Software emulation<br>• Archive users with active software licences<br>• Software vendors and/or patent holders releasing previously proprietary formats as open standards (as occurred with .mp3) | Aliza Leventhal, *Designing the Future Landscape: Digital Architecture, Design & Engineering Assets,* A report on the Architecture, Design and Engineering Summit organized by the Library of Congress, the National Gallery of Art and the Architect of the Capitol on November 16 & 17, 2017 at the Library of Congress: http://digitalpreservation.gov/meetings/ade/ade2017.html accessed 20 June 2018<br><br>David S. H. Rosenthal, Emulation & Virtualization as Preservation Strategies, Andrew W. Mellon Foundation, 2015: https://mellon.org/Rosenthal-Emulation-2015/ accessed 20 July 2018<br><br>Autodesk Inc. offers a number of freely-downloadable file viewers from its website, enabling access to various proprietary Autodesk file formats: https://www.autodesk.com/viewers/all-viewers accessed 18 June 2018<br><br>Autodesk Inc. also offers a free, online (browser-based) file viewer that claims compatibility with over 50 file formats: https://viewer.autodesk.com/ accessed 18 June 2018<br><br>Robert McNeel & Associates currently offer downloadable evaluation versions of widely-used Rhinoceros 3D modelling software (versions 5 and 6) on their website, which continue to function for viewing and editing files after the end of the 90-day trial period (sans plugins and save/export functionality) thus enabling long-term access to the proprietary .3dm file format. Rhinoceros is also able to import and parse (interpret) a wide variety of other proprietary file formats. https://www.rhino3d.com/download accessed 18 June 2018 |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| 3.7 Retain functioning legacy hardware, especially physical carrying media drives (e.g. tape drives, floppy and zip disk drives etc.) for as long as they are required for accessing and transferring files stored on physical carrying media.<br><br>When no longer required, offer functioning legacy hardware as a donation to a collecting institution. | Legacy hardware (especially mechanical hardware with moving parts and/or rubber belts that may degrade over time, for e.g. tape drives) is rapidly disappearing as organisations dispose of apparently obsolete technology *en masse*. Functioning tape and disk drives for early media formats are particularly rare and may be valuable for other organisations looking to recover files stored on legacy media. | The Computer Archaeology Laboratory at Flinders University is an Adelaide-based organisation that collects and preserves functioning legacy hardware and actively and preserves legacy software. The Laboratory accepts donations: http://csem.flinders.edu.au/thegoodstuff/comparch/about.php accessed 18 June 2018<br><br>The Monash Museum of Computing History is based in Melbourne: https://www.monash.edu/it/about-us/museum-of-computing-history/about-the-museum accessed 18 June 2018<br><br>The Australian Computer Museum Society Inc. is based in Sydney: https://www.acms.org.au/ accessed 18 June 2018 |
| *4. Naming and filing procedures* | | |
| 4.1 Practices should adopt a consistent and well-documented file naming convention.<br><br>At minimum file names should include a short specific plain English title, a date (possibly as a six or eight-digit code e.g. 27062018) and a variant or version number (e.g. v1_1 means variant 1 version 1).<br><br>Avoid 'illegal' characters and symbols in file names.<br><br>Use folder names to assign context to groups of files. | Architectural practices 'need to organise their IT filing system in a much stricter way than their paper records have been organised' (Peycerê 2009 p. 4).<br><br>Accurate file naming, including a date and variant or version number, makes it easier to find files (especially the latest version files). This is especially important when working in a collaborative environment.<br><br>Good file naming practice is a habit that can and should be learned.<br><br>'Even if every practice has to define its own rules about managing its records, these rules, if applied consistently within each practice, should help us archivists a lot in our task of accepting and managing an archive' (Peycerê 2009 p. 2) | See https://www.mtu.edu/umc/services/digital/writing/characters-avoid/ for characters to avoid in file names. |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| 4.2 Consider implementing a quality management system guided by ISO 9000; certification need not necessarily be the final goal. | '…the development of a quality management system and ISO 9000 certification to guarantee its implementation, are a real help to architectural practices in structuring their data and creating a solid, reassuring archive structure. The ISO 9000 standard can be used as an inspiration without being fully implemented and without certification necessarily being the final goal' (Peyceré 2009 p. 4). | ISO 9000 family – Quality Management https://www.iso.org/iso-9001-quality-management.html accessed 27 June 2018 |
| *5. Retention, disposal and donation schedules* | | |
| 5.1 Architectural practices should develop a policy documenting the types of records (analogue and digital) that will be archived, as well as retention and disposal schedules for the records in their care.<br>This policy should be revisited and revised on a scheduled basis (e.g. once every five years).<br><br>Record and retain a list of what has been destroyed (paper records) or deleted (digital records)<br><br>5.2 While developing retention and disposal schedules, architectural practices should address the question of donating records to collecting institutions (libraries, archives or museums).<br><br>Intentions to donate records (which records will be donated, when and any conditions on public access) | Traditionally, architectural practices have determined their own retention and disposal schedules for records in their care, sometimes on an ad-hoc basis.<br>Once destroyed or deleted, records are lost forever, and in many cases, a decision to destroy or delete records is later regretted.<br>Transparent retention and disposal schedules should be based on available guidelines and will reduce the likelihood that useful, valuable or historically significant records will be destroyed or deleted.<br><br>From the UNESCO Guidelines for the Preservation of Digital Heritage: 'decide what materials definitely must be preserved and for how long; what definitely does not need to be preserved; and what should be accepted for interim preservation action while a more definitive selection decision can be made' (p. 72).<br><br>Factors to be considered in determining schedules for retention and/or disposal of records include:<br>• Legal requirements<br>• Utility of records for reuse<br>• Promotional value<br>• Historical significance | David Peyceré, "The Architectural Practices as First Curators of their Archives" NAi, Rotterdam. Hybrid Architectural Archives Conference (11-12 June 2009), conference.nai.nl/mmbase/attachments/525844/1.1_David_Peyceré.pdf accessed 14 June 2018<br><br>National Library of Australia, *Guidelines for the Preservation of Digital Heritage,* UNESCO (2003) http://www.unesco.org/new/en/communication-and-information/resources/publications-and-communication-materials/publications/full-list/guidelines-for-the-preservation-of-digital-heritage/ accessed 14 June 2016<br><br>Waverly Lowell and Tawny Ryan Nelb's *Architectural Records: managing design construction records* (Chicago: Society of American Archivists, 2006) discusses the appraisal of paper-based architectural records for retention or disposal (pp. 69-85) and provides a table (pp. 84-85) suggesting retention, disposal and reformatting actions for a variety of record types. While aimed primarily at archivists and analogue (paper and physical model-based) media, it provides an indication of the kinds of records that collecting institutions may consider worthy of long-term preservation. |

| Key Guidelines | Rationale | Resources and further reading |
| --- | --- | --- |
| should be documented in schedules of retention and disposal. | | More research is needed on the kinds of cultural information that may be encoded in digital architectural records; however, architects, with a working knowledge of the special affordances of 3D CAD and BIM models (e.g. parametric relationships and modelling history that may reveal design intent etc.) are well placed, using the above paper as a starting point, for assessing the cultural value of digital architectural models and drawings in their custody.<br><br>Recent research projects utilising the information encoded in native 3D/BIM models are mentioned in Aliza Leventhal, *Designing the Future Landscape: Digital Architecture, Design & Engineering Assets,* a report on the Architecture, Design and Engineering Summit organized by the Library of Congress, the National Gallery of Art and the Architect of the Capitol on November 16 & 17, 2017 at the Library of Congress: http://digitalpreservation.gov/meetings/ade/ade2017.html accessed 20 June 2018<br><br>The Canadian Centre for Architecture's (CCA) *Archaeology of the Digital* exhibition and book (Montreal: Canadian Centre for Architecture and Sternberg Press, 2013), curated and edited respectively by Greg Lynn, investigates the history of computing in architecture using evidence provided, in part, through the interrogation of digital architectural records: https://www.cca.qc.ca/en/events/3333/archaeology-of-the-digital accessed 18 June 2018 |
| 5.3 While architects have traditionally donated records near the ends of their careers, donating digital architectural records should be donated earlier rather than later. | Digital records are more easily destroyed through benign neglect than paper records. Donating digital records earlier to a collecting institution with the technological capabilities to preserve digital files long-term improves the likelihood that records will pass successfully into long-term preservation.<br><br>In addition, 'Selection and transfer of data is most likely to be successful when archivists can consult with records creators' (Walsh 2015 p. 17). If files are transferred early, 'records creators are more likely to remember and share crucial information about their creation, modification, use, and organisation, allowing archivists to record more thorough and accurate contextual information for users of the archive' (p. 17). | |
| 5.4 Exercise caution when disposing of records, especially digital records.<br><br>If practices intend to donate records to collecting institutions in the future, err on the side of retaining more rather than less.<br><br>"If in doubt, don't throw it out!" | Given that 'a decision not to preserve is usually a final one for digital materials,' the UNESCO *Guidelines for the Preservation of Digital Heritage* recommend a cautious approach to the disposal of digital records (p. 72).<br><br>The potential historical significance of architectural records may not be immediately obvious, and the future research value of digital architectural records (3D CAD and BIM models in particular) is difficult to predict.<br><br>All categories of records are likely to be of interest to future historians, including but not limited to, correspondence (emails, faxes), sketches and design | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| | development drawings and models, design intent and as-built CAD models and documentation, photographs, physical models, renderings, animations, presentations to clients (i.e. PowerPoint), presentation drawings, specifications, reports and contracts.<br><br>‘The universe of users of design and construction records is diverse. It includes building owners, architects, preservationists, historians, planners, students, scholars, engineers, lawyers, model builders, horticulturalists, manufacturers, art curators, local communities, and others. Uses include restoration, adaptation and additional construction, historical research, litigation, analysis of sociological factors and patterns of use, publication, exhibition, licencing of designs, educational interpretation, inspiration for new design, the study of drawing techniques, and seismic renovation among others’ (Lowell & Nelb 2006 p. 70). | |
| 5.5 As a minimum, practices should retain the following records for each project: 1) milestone project deliverables (outlined in more detail in 5.6), including as-built and design intent contract documentation and specifications; 2) a selection of design development records (including for e.g. sketches and studies); 3) photographs, especially construction progress photographs. | Lowell & Nelb state that ‘the most significant project records are the design process and final construction records, and photographs. These records are required for understanding the development of the program’ (p. 73).<br>Sketches and studies ‘document the creative process’ while as-built or record drawings ‘are the most accurate documentation of the project at its completion’ (p. 43). Meanwhile, construction progress photographs may record changes that have occurred during construction when as-built project record drawings were not created by the contractor. | |
| 5.6 Retain at least a minimum of records for unbuilt projects. | ‘Unbuilt project records play an important role in understanding a designers’ overall work and may be significant from a design perspective for historical research’ (Lowell & Nelb 2006 p. 74). | |

| Key Guidelines | Rationale | Resources and further reading |
|---|---|---|
| 5.7 Set aside time at the end of a project for deliberately organising and archiving project records. | | ISO 15489 Information and documentation – Records management https://www.iso.org/standard/62542.html accessed 27 June 2018 |
| 5.7 As outlined in the *NATSPEC National BIM Guide,* milestone digital project deliverables are to be submitted to clients and other stakeholders on CD/DVD recordable media with data clearly organised and software versions labelled.<br><br>Digital deliverables for BIM workflows are outlined in the *NATSPEC National BIM Guide* (pp. 19-20).<br><br>For non-BIM workflows, in lieu of BIM construction coordination and design intent models, digital deliverables should include final versions of all 2D and 3D CAD files in both native CAD formats and at least one open format.<br><br>*Best practice:* submit milestone digital project deliverables to clients and other stakeholders on M-DISC CD/DVD recordable media. Milestone digital deliverables should also be ingested into the central archival repository in a labelled folder with data clearly organised and software versions labelled. | This practice will ensure that milestone project files (at the very least) are captured in a 'preservation ready' open-source format, along with their associated native format 3D CAD/BIM files.<br><br>M-DISC CD/DVD recordable media potentially offers a longer archival lifespan than conventional dye-based optical media. However, M-DISC is still an unproven technology over long timescales; meanwhile, the small capacity of M-DISC (4.7GB) makes it unsuitable for archiving large quantities of data. | NATSPEC, *NATSPEC National BIM Guide*, Construction and Information Systems Limited (2016) https://bim.natspec.org/documents/natspec-national-bim-guide accessed 14 June 2018<br><br>The archival properties of M-DISC media and a rationale for use are discussed in the literature review contained within the *Archiving digital architectural records: towards a national framework* report, and in a report by Ivan Svrcek, *Accelerated Life Cycle Comparison of Millenniata Archival DVD,* China Lake, California: Life Cycle and Environmental Engineering Branch Naval Air Warfare Center Weapons Division (2013), https://archive.org/details/ChinaLakeFullReport accessed 14 June 2018 |

## REFERENCES


Ball, Alex *Preserving Computer-Aided Design (CAD),* DPC Technology Watch Report, Bath, UK: Digital Preservation Coalition (2013), http://dx.doi.org/10.7207/twr13-02 accessed 14 June 2018

Barrera-Gomez, Julianna and Ricky Erway, *Walk This Way: detailed steps for transferring born-digital content from media you can read in-house*, Dublin, Ohio: OCLC Online Computer Library Centre, Inc. (2013), http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf accessed 14 June 2018

Collins, Julie, Susan Collins and Christine Garnaut, "Behind the Image: assessing architectural drawings as cultural records" in *Archives and Manuscripts,* Vol. 35 No. 2 (November 2007) pp. 87-107

Digital Preservation Coalition, *Digital Preservation Handbook (2nd Edition)* (2015), https://www.dpconline.org/handbook accessed 27 June 2018

Dollar, Charles and Lori Ashley, *Digital Preservation Capability Maturity Model (DPCMM): Background and Performance Metrics Version 2.7* (July 2015), https://www.statearchivists.org/resource-center/resource-library/digital-preservation-capability-maturity-model-dpcmm/ accessed 14 June 2018

Leventhal, Aliza *Designing the Future Landscape: Digital Architecture, Design & Engineering Assets,* A report on the Architecture, Design and Engineering Summit organized by the Library of Congress, the National Gallery of Art and the Architect of the Capitol on November 16 & 17, 2017 at the Library of Congress: http://digitalpreservation.gov/meetings/ade/ade2017.html accessed 20 June 2018

Lowell, Waverly and Tawny Ryan Nelb, *Architectural Records: managing design construction records,* Chicago, Ill: Society of American Archivists (2006)

National Archives, The, (UK) *Preserving Digital Records: Guidance* http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/guidance/ accessed 27 June 2018

National Library of Australia (NLA), "Mediapedia" https://mediapedia.nla.gov.au/home.php accessed 14 June 2018
National Library of Australia, *Guidelines for the Preservation of Digital Heritage,* UNESCO (2003), http://www.unesco.org/new/en/communication-and-information/resources/publications-and-communication-materials/publications/full-list/guidelines-for-the-preservation-of-digital-heritage/ accessed 14 June 2018

NATSPEC, *NATSPEC National BIM Guide*, Construction and Information Systems Limited (2016) https://bim.natspec.org/documents/natspec-national-bim-guide accessed 14 June 2018

Peyceré, David "The Architectural Practices as First Curators of their Archives" NAi, Rotterdam. Hybrid Architectural Archives Conference (11-12 June 2009), conference.nai.nl/mmbase/attachments/525844/1.1_David_Peyceré.pdf accessed 14 June 2018

Phillips, Megan, Jefferson Bailey, Andrea Goethals and Trevor Owens, *The NDSA Levels of Digital Preservation: An explanation and uses,* Library of Congress, National Digital Stewardship Alliance (2013), www.digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf accessed 16 May 2018

RLG/OCLC Working Group on Digital Archive Attributes, *Trustworthy Digital Repositories: Attributes and Responsibilities,* Mountain View, California: RLG (May 2002), www.digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf accessed 14 June 2018

Rosenthal, David S. H. Emulation & Virtualization as Preservation Strategies, Andrew W. Mellon Foundation, 2015: https://mellon.org/Rosenthal-Emulation-2015/ accessed 20 July 2018

Rosenthal, David S. H. Thomas Robertson, Tom Lipkisi, Vicky Reich and Seth Morabito, "Requirements for Digital Preservation Systems: A Bottom-Up Approach" in *D-Lib Magazine,* Vol. 11 No. 11 (November 2005), www.dlib.org/dlib/november05/rosenthal/11rosenthal.html accessed 14 June 2018

Svrcek, Ivan, *Accelerated Life Cycle Comparison of Millenniata Archival DVD,* China Lake, California: Life Cycle and Environmental Engineering Branch Naval Air Warfare Center Weapons Division (2013), https://archive.org/details/ChinaLakeFullReport accessed 14 June 2018

Zaluendo, Ines Maria, ‘Paradigm Shift: curatorial views on collecting and archiving architectural drawings in an evolving born-digital landscape’ paper presented at the Society of American Archivists Conference, Washington D.C. (August 2014) p. 3, https://dash.harvard.edu/bitstream/handle/1/13442962/IZ_Born-Digital%20Architectural%20Drawings_FINAL_2014_A1b.pdf accessed 17 May 2018

## 6. Directions for future research and for refinement and presentation of the framework

Several directions for future research, and for further development, refinement and presentation of the draft national framework for archiving digital architectural records documents are suggested:

- The framework has been tested with a limited number of stakeholders in architectural practice and the archival profession, each of whom read the document and provided feedback and suggestions based on their professional expertise. The framework could be tested and refined through obtaining more extensive feedback from selected representatives of the two user groups. Two approaches are recommended:

  (1) practical testing through implementation of the guidelines in architectural practice and collecting institutions over a period of 3-6 months.

  (2) a series of workshops with architectural practitioners and archival professionals to introduce, explain, review and discuss the themes, guidelines and practical aspects of the framework and to obtain feedback from stakeholders.

- The framework content could be substantially enhanced through improved graphic presentation, potentially bringing the document into visual or brand conformity with other NATSPEC documents, including the *NATSPEC National BIM Guide*.
- Practical recommendations could be further developed and disseminated to practitioners and professionals through a series of short interactive or video tutorials.
- Likewise, decision making processes (for example those related to digital records retention and disposal) and practical procedures (for example those related to ingest of data on physical carrying media) could be codified in a series of graphic flowcharts.