

EMULATING THE PAST, ENVISIONING THE FUTURE: A STUDY OF CAD FILE PRESERVATION IN AUSTRALIAN ARCHITECTURAL PRACTICES

Kirsten Day¹

*Faculty of Architecture, Building and Planning, University of Melbourne, The University of Melbourne,
Parkville, Victoria, 3010, Australia*

In the 1980s; computer-aided design (CAD) fundamentally changed how architectural documentation was produced; moving from analogue to digital by the early 2000s. This study extends existing preservation work; examining the archival difficulties we encounter with CAD data, such as unreliable media carriers; software supersession; and obsolescence; alongside emulation potential for historical and modern files. Emulation facilitates access to original CAD files and associated metadata, which derivative formats like PDFs do not retain. The study contrasts early CAD files' simplicity with the advanced interoperability of contemporary files, which introduces significant emulation challenges. Moreover, rapid software updates aimed at 'innovation' often disregard the long-term accessibility of files beyond the three-year migration window. The paper stresses the importance of maintaining CAD files in their original environment to preserve the creator's intent and calls for proactive strategies to maintain the integrity of architectural documentation.

Keywords: CAD preservation; digital archiving; emulation; architectural documentation; software lifecycle management

INTRODUCTION

Architectural, Engineering, and Construction Management (AECM) consultants create many different documents for building projects. These might include feasibility studies, concept design, design development, construction drawings, and virtual models for managing building information. Nowadays, all these documents originate in digital form. Moreover, keeping records for specified durations is necessary for quality control, legal, and liability reasons. For example, the Australian Institute of Architects recommends keeping records for 10 years (Australian Institute of Architects 2022). Many of these documents are perfunctory and can be stored in Portable Document Format (PDF), a static format that is easy to view in multiple software applications. However, older files may be difficult to access because they were created on obsolete software or stored on outdated media carriers (i.e. floppy disk, zip disk or CDs).

Software obsolescence

In a contemporary setting, AECM consultants communicate the design and assembly of a building using specialised software (programs and operational systems). The

¹ dayk@unimelb.edu.au

most prominent is Computer-Aided Design (CAD), although many other software applications are used in offices for graphics, rendering, communication, and office management. Today, CAD is provided via a subscription service, with the most prominent global provider being Autodesk. This business model focuses on front-end innovation and interoperability but has limited support for archival retrieval (Autodesk Support 2023b) The process for accessing older files is to manually migrate data incrementally across different software versions—eventually opening these files in the latest version of whichever software product is required (Autodesk Support 2023a). Subscriptions are tracked using IP Addresses. This is a very different story from the 1990s when pirated or 'cracked' copies of software could not be traced. The definition of an old or obsolete file is four versions—usually produced over three years. Access to significantly older documents—say, in the case of renovation or retrofit—will be more complicated.

Many AECM practices that operated through the 1990s and 2000s experimented with many different types of software—thereby, many archives include the enormous variety of early CAD software from companies that no longer exist and/or files that have been saved in unique formats—meaning that other software programs cannot open them (i.e. not saved in generic formats such as DWG). Significantly, the industry has moved from 2D drawings to 3D modelling, and with this shift, there is an increased complexity of files and any associated links.

For preservationists working with files, data degradation (or bit-rot) is a well-documented risk with digital records and is fundamental to preservationists' work. Of this, Leventhal and Thompson highlight three significant risks for born-digital records, including the availability of the original software versions that produced the file, missing support files, accompanying metadata, and/or broken or lost links to external reference files when born-digital design records are transferred from a record creator to an archival institution, and understanding how the nuances of the iterative design processes are reflected in a project's digital born design files, including the multiple uses of design software and varying practices of design professional or a firm across projects (Leventhal and Thompson 2021, 6-7).

Media carrier obsolescence

Another barrier to accessing digital media is the medium on which data is stored. Digital archival methods have evolved significantly, transitioning from magnetic tapes, CDs, and hard drives to more advanced options. In the early 1990s, the high-density 3 1/2-inch floppy disk was the standard for storing 1.44MB—a significant improvement from the late 1980s. Today, contemporary storage media, such as thumb drives, offer capacities ranging from gigabytes to terabytes, and the cloud provides petabyte-scale storage. However, many older storage formats, like the Zip drive, have become obsolete.

Retrieving data from these outdated sources may require discontinued hardware and software that is often difficult to find. Those searching for functioning hardware or compatible software typically consult niche forums, enthusiast communities, or specific Reddit threads. Further, as Swalwell describes, "the physical media on which legacy digital objects are stored—whether magnetic or optical (i.e. floppy disks, CD-ROMS)—deteriorate over time" (Swalwell 2022).

Preservation and Emulation as a Service

A possible solution to these concerns is to create disk images of obsolete media carriers, which can then be run under emulation. Several institutions have explored

emulation, including the Canadian Centre for Architecture (Yaneva 2020; Allen 2022; De Vletter 2019) and Harvard Graduate School of Design (See Whiteside *et al.*, 2021). A potential solution to access is 'emulation as a service infrastructure,' often referred to as EaaS, which allows users to view original files in their native environment. This extends the utility of digital documents beyond the lifecycle of the software with which they were created. Emulation practices are computation and technical processes that allow one system to reproduce the functions and results of another (Acker 2021, 1148). The direction of this research emerged from an Australian Research Council (ARC) Linkage Infrastructure, Equipment and Facilities (LIEF) project—The Australian Emulation Network: Born Digital Cultural Collections Access. This has initiated the AusEaaS network and is now in the process of Part 2 funding to firmly establish this initiative. Part of this project is the working group for Architecture and Design that specifically explores architectural and industrial design, interrogating how emulation can be an asset not only to archival and preservation concerns—but to the everyday practice of AECM companies.

Scholarly Debate

The research concerning digital data in architecture is divided into two primary areas. The first area emphasizes the development of standardised record-keeping practices. As highlighted by NATSPEC's work (See Burns 2018), there is an effort to establish record-keeping protocols consistent with ISO 9001 standards, which are vital for AECM quality control management. Record-keeping practices often convert files to generic formats, which may simplify complex environments. Although this is functional for maintaining records, it may not be ideal for the continuous development of projects or for accessing data required for building renovations or retrofits.

The second area of research focuses on preserving and archiving digital records. Archivists and institutions, including libraries and architectural museums, have developed comprehensive policies for collecting and maintaining architectural documentation (Ball 2013; Owens 2018; De Vletter 2019; Whiteside *et al.*, 2021; Leventhal and Thompson 2021; Winn 2022). Emulation research is gaining traction in the preservation field, aiming to incorporate this technology into Gallery, Library, Archive, and Museum (GLAM) interfaces to enhance access to archived digital materials (Acker 2021; Allen 2022; Acker 2021). Emulation is an emerging field within digital preservation; however, it faces several challenges, including issues related to using copyrighted CAD software outside of research institutions. The suitability of emulation as a long-term solution for the industry remains an open question, and this uncertainty serves as a starting point for the broader investigative aspects of this study.

The Claim of Significance and the Author's Argument

The current business model for modern AECM software prioritises immediate innovation over preserving long-term access to digital files. This software is part of a continually evolving ecosystem of interoperable software. Regular software updates are required, usually on an annual basis, to ensure smooth operation. However, these updates can disrupt links and lead to the loss of metadata unless special care is taken during file migration, which is currently a manual process.

From an archival standpoint, significant efforts are made to access and maintain files stored on outdated media or created with obsolete software—efforts illuminating the challenges we face in preserving the history of the late 20th century. Using emulation to access these files, especially with early Computer-Aided Design (CAD) programs

like AutoCAD, is becoming increasingly complex, especially for simple two-dimensional files.

This paper contends that the focus on innovation in CAD software has overlooked the essential requirement for enduring access and management of digital design documents. Today's business strategies do not sufficiently account for the complete lifecycle of building management, from inception to demolition. While emulation has been used to access older files, it falters with more intricate, contemporary files—particularly three-dimensional models—which exceed the capabilities of traditional two-dimensional archival techniques. AECM practitioners need a comprehensive approach to digital file management that encompasses the whole lifespan of a building, challenging the inadequacy of the current standard of a three-year time frame to accommodate the numerous phases of a building's life - and the ongoing interaction of design professionals with the building.

This paper advocates for the urgent need for further digital preservation research, specifically focusing on utilising emulation to access legacy files. Emulation technology, though relatively nascent, holds significant potential for the broader application of digital-born materials in archival preservation. However, its critical importance to the AECM sector's ongoing data storage and access demands thorough investigation.

The rapid pace of front-end innovation exacerbates the challenges of accessing digital files. Software upgrades and migrations frequently result in the loss of valuable metadata or necessitate saving files in generic formats like PDFs, undermining the original purpose of complex files such as 3D models. Furthermore, there is an urgent need to address the proper data management and maintenance practices once an AECM project is considered complete or the budget is exhausted.

This study highlights these pressing issues, advocating for a comprehensive exploration of emulation's potential role and benefits. The goal is to support robust, long-term archival strategies within the AECM industry, ensuring the preservation and accessibility of digital assets for future generations.

METHOD

This research project's approach was structured in two distinct stages. The first stage conducted an exploratory case study on the emulation of CAD files from the archive of Australian Architect Gregory Burgess (GBA). This was part of the ARC Grant "Australian Emulation Network: Born Digital Cultural Collections Access". This distinct piece of work focused on exploring early CAD—particularly the exploratory work of the 1990s when there were many more software developers/providers (i.e. CADKEY, MicroStation, Personal Designer, etc.). The goal was to document the emulation process needed to faithfully recreate the software environment initially used to create these CAD files. Results from this investigation were presented at the Born Digital Cultural Heritage Symposium (BDCH#23) in December 2023, speculating on the complexity of emulating more sophisticated 3D files that interact with other software systems, such as contemporary Building Information Modelling (BIM) files.

Building on those initial findings, the second stage involved interviewing industry professionals. The intent was to understand the broader implications of the case study results and gain insight into current industry practices concerning software usage and strategies for saving and archiving project data. The Ethics Approval, supporting and summary documentation can be reviewed at DOI: 10.26188/26042842.

Table 1: Participant details

	Practice size (pax)	Australia / New Zealand offices (no)	International offices
Design practice 1	11,000	43	124
Design practice 2	650	3	5
Design practice 3	180	4	0
Design practice 4	165	4	0

Open-ended questions were conducted via Zoom interviews with practice managers, IT managers, and senior designers from four mid- to large-sized firms in Australia with multiple offices that incorporate national and international operations. Although this part of the research had a broader scope, it confirmed significant concerns regarding data accessibility, the practical demands of software management, and apprehensions related to file preservation.

FINDINGS

Software

Emulation: Access and Copyright

The copyright terms set by software providers restrict access to obsolete software. ARM Architects generously provided full licenses for AutoCAD14 and AutoCAD 2000. However, Autodesk's business practices limit opening files created before 2020. These files are only accessible if you are in the process of migrating to the current software version (Autodesk Support 2023b). Direct inquiries to Autodesk yielded no workaround; we received an apology for any inconvenience from a well-meaning bot.

Copyright laws must be navigated to access these files legally. Swalwell (2022) notes that recent amendments to the Copyright Act—specifically the "Research Exception" found in Division 3: Libraries and Archives 113J—allow libraries and archives (as defined by the Act) to create copies of copyrighted materials for research purposes (Australian Government 2024). These copies can then be used within the library or archive or be shared with other such institutions, provided specific criteria are met. In Australia, this exception facilitates the use of emulation for research purposes, with legal frameworks varying across different jurisdictions.

Interviews: Migration of data, stewardship, and intellectual property

Data migration is a time-consuming manual process, and with the vast spectrum of data collected by AECM practices, determining what needs to be retained and how this is retained requires further investigation. While this can be managed during a project's design and construction lifecycle, file management becomes unclear after the contract has been concluded. Stewardship and ongoing maintenance to ensure ongoing access to files is often not specified. In a more traditional contract where the architect is the lead consultant—it could be assumed that it is a part of the architect's archive. However, this role becomes murky when there is no contractual arrangement for ongoing maintenance and guidelines about the length of time records are kept (10 years). Many architects interviewed described being asked for documents 15-20 years after the conclusion of a design project. This period now includes 3D and BIM models.

The issue of stewardship and intellectual property (IP) has become increasingly complex in contemporary software that supports collaborative design work. In the

Australian construction industry, where Design and Construct (D&C) and Novated Design and Construct (NDC) project models are standard, the contractor assumes responsibility and risk for the entirety of the project in the case of D&C or post-Novation in the case of NDC contracts. However, the ownership and management of the project documentation are often undefined. Intellectual property considerations further complicate this ambiguity over documentation control, which can result in reluctance to share copies of files even for record-keeping purposes.

Storage

Emulation

In our research on archives, we have dealt with various data storage media, ranging from cassettes, floppy disks, and Zip drives to CDs, external hard drives, and cloud storage. Large practices often use cloud services such as AWS, to which access is provided via third parties and introduces risk. Business model changes or management shifts can jeopardise data access. Although AWS is robust, barring catastrophic events, the vulnerability often lies with these third-party gatekeepers (such as Salesforce), whose policies might leave architectural firms in precarious situations.

Over time, storage media like CDs have proven volatile. The early disks can be challenging to access not due to storage failure but because the playback equipment, like Zip disk readers, is no longer manufactured, making access sporadic or dependent on specialised services. Emulation has been instrumental in accessing such data, yet this approach relies on copyright permission within research institutions—a privilege currently accessible in Australia but not in many other places. This presents another consideration that needs addressing in digital data preservation.

Interviews: The cloud and futureproofing

During the interviews, an unexpected issue surfaced: reliance on cloud storage. While fundamental to businesses operating offices in multiple locations and jurisdictions—many of these methods are 'opaque' in operation. Some interviewees cited "techno-feudalism" as a necessary evil (see Varoufakis 2023). The conversations covered various topics, including the types of media used for file storage, the increasing size of files, and the strategies for managing and updating archives to maintain access. Many practices have a mix of analogue and digital data, and there were frequent mentions of files saved on now obsolete media like zip disks and CDs. The participants expressed surprise regarding the instability of these storage media, particularly optical media like CDs.

A key point of discussion was the universal shift towards cloud storage, with most, if not all, practices using services like AWS. Interestingly, access to AWS often relies on third parties that are not immune to failure. Concerns were raised about the potential impacts of changes in business plans, ownership, or strategies, questioning the continued access to data in such events.

While some practices have adopted multiple methods for saving files, this is not widespread. A few practices maintain an in-house back up system, with daily backups stored on-site and weekly backups off-site. However, this is more of an exception than the norm.

The conversation also touched upon quantum computing and its implications for our capacity to store ever-increasing amounts of data. Considerations about the sustainability of storing vast quantities of data were raised. It was noted that

sometimes, it is more economical to build physical archives and to print documents than to rely on cloud storage. It was concluded that this topic warrants a separate, in-depth discussion at another time.

CONCLUSION

As the construction industry marches on with its rapid front-end innovation, project documentation becomes increasingly complex, demanding a new approach to ensuring the long-term accessibility of digital records. Drawing a parallel with Amazon's evolution from a simple bookseller to a complex web services provider, Autodesk has similarly transformed. It now extends beyond traditional CAD software to offer web services through third-party platforms, reflecting a broader trend of diversification and integration within the industry.

Managing digital storage, particularly the phenomenon known as "Cloud Sprawl"—the unchecked proliferation of cloud services—presents a new set of challenges. This expansion underscores the urgent need for standardised processes and terminology that bridge the gap across professions. IT experts, AECM specialists, archivists, and preservationists must collaborate to establish what data should be preserved, ensuring the retention of valuable links and metadata.

Interviews with industry professionals reveal a dichotomy between the allure of innovation and the fundamental need for data preservation. The emergence of hybrid archives and the current software duopoly present challenges not seen since the software diversity of the 1990s, when many software programs became obsolete, and the Mac vs. PC era promoted incompatibility not only between operating systems, software, and media storage. While this still exists today—it is not as pervasive. The issue of technological obsolescence further complicates matters. The dwindling availability of compatible devices and the enduring necessity of physical connections in an increasingly wireless world underscore the importance of maintaining older technology to uphold digital legacies.

Moreover, the risk of digital data becoming inaccessible looms large without a consistent strategy to guarantee longevity. Despite existing guidelines on digital preservation, there's a conspicuous lack of strategic foresight in safeguarding the future of digital archives—including data stewardship—a pressing concern in the construction industry where the extended use of data is indispensable for building maintenance, retrofitting, or material recycling.

The necessity for enduring access to digital data in the AECM field cannot be overstated. With large files and complex interdependencies, the need for a standardised preservation system is apparent. This paper has highlighted Autodesk's role in this dynamic, pinpointing the widening gap between innovation and the ability to preserve and access digital data over time.

Emulation, while applicable in archival and preservation contexts, currently falls short of AECM needs due to its limitations in interacting with files and the current restriction of access to research institutions. As file sizes increase and become more complex, the fragility of data also grows during software migration. AECM companies must advocate for better document access, and if software providers do not adapt, alternative solutions may need to be sought. The industry's demand for improved methods is crucial as emulation, in its current form, is insufficient for the multifaceted requirements of AECM documentation and data use.

ACKNOWLEDGEMENTS

The author extends gratitude to ARM Architects for their generous contribution of software licenses. This research was conducted under the auspices of an Australian Research Council LIEF Grant, with the author serving as a principal investigator. The data associated with this project will be made available at DOI: 10.26188/26042842

REFERENCES

- Acker, A (2021) Emulation practices for software preservation in libraries, archives and museums, *Journal of the Association for Information Science and Technology*, **72**(9), 1148-60.
- Allen, M (2022) SCOPE and EaaSI: Software emulation and the future of architectural history, *Journal of the Society of Architectural Historians*, **81**(1), 122-23.
- Australian Government (2024) *Copyright Act 1968 Vol C2024c00092 (C62)*, Available from: <https://www.legislation.gov.au/C1968A00063/latest/text> [Accessed 15 July 2024].
- Australian Institute of Architects (2022) *Document Retention and Destruction*, Acumen Practice Notes 30 November 2022, Available from: <https://acumen.architecture.com.au/practice/office-administration/archiving/document-retention/> [Accessed 15 July 2024].
- Autodesk Support (2023a) *Compatibility between Older Versions of AutoCAD and Newest Versions*, AUTODESK, Available from: <https://www.autodesk.com/support/technical/article/caas/sfdcarticles/sfdcarticles/Do-you-need-to-uninstall-older-versions-of-AutoCAD-before-installing-a-new-version.html#:~:text=A%20subscription%20to%20Autodesk%20software,than%20one%20version%20of%20AutoCAD> [Accessed 8 October 2023].
- Autodesk Support (2023b) Which Are the Officially Supported Product Versions of Revit?, AUTODESK, Available from: <https://www.autodesk.com/support/technical/article/caas/sfdcarticles/sfdcarticles/Which-are-the-officially-supported-product-versions-of-Revit.html#:~:text=To%20keep%20the%20efforts%20of,%2C%202022%2C%202021%2C%202020> [Accessed 8 October 2023]
- Ball, A (2013) *Preserving Computer-Aided Design (CAD)*, 1st Edition, Bath, UK: Digital Preservation Coalition.
- Burns, C (2018) *Archiving Digital Architectural Records: Towards a National Framework*, Adelaide, Australia: University of South Australia Architecture Museum and NATSPEC Construction Information.
- De Vletter, M (2019) *Don't Be Afraid of the Digital*, *Arts*, **8**(1), 6.
- Leventhal, A and Thompson, J (2021) *Preserving Born-Digital Design and Construction Records*, Bath, UK: Digital Preservation Coalition.
- Owens, T (2018) *The Theory and Craft of Digital Preservation*, Baltimore: Johns Hopkins University Press.
- Sayegh, E (2020) *The Dangers of Cloud Sprawl: When There's Too Much Cloud Going on*, Forbes 28 April, Available from: <https://www.forbes.com/sites/emilsayegh/2020/04/28/the-dangers-of-cloud-sprawl-when-theres-too-much-cloud-going-on/?sh=311790971852> [Accessed 15 July 2024].
- Swalwell, M (2022) The Australian Emulation Network: Accessing Born Digital Cultural Collections, In: Proceedings of ISEA2022, *International Symposium on Electronic Art (Proceedings of the Second Summit on New Media Art Archiving)*, Barcelona, 10-16 June.

- Varoufakis, Y (2023) *Techno-Feudalism: The Demise of Capitalism and the Rise of a New Power*, London: Penguin Random House.
- Whiteside, A, Abrams, S, Rogers, S, Kao, K and Kataw, H (2021) *Building for Tomorrow*, Institute of Museum and Library Services, Available from:
https://projects.iq.harvard.edu/files/buildingtomorrow/files/building_for_tomorrow_w_hitepaper_version_1.0.pdf [Accessed 15 July 2024].
- Winn, S (Ed) (2022) *Born-Digital Design Records Trends in Archives Practice*, Chicago: Society of American Archivists.
- Yaneva, A (2020) *Crafting History: Archiving and the Quest for Architectural Legacy Expertise: Cultures and Technologies of Knowledge*, Ithaca: Cornell University Press.