EVALUATING DIGITAL PRESERVATION CAPABILITY WITH LARGE AT-RISK COLLECTIONS

Lessons learnt from preserving the NVA Archive

Leo Konstantelos
University of Glasgow
United Kingdom
Leo.konstantelos@glasgow.ac.uk

Clare Paterson
University of Glasgow
United Kingdom
Clare.Paterson@glasgow.ac.uk

Emma Yan
University of Glasgow
United Kingdom
Emma.Yan@glasgow.ac.uk

Abstract – This paper presents the efforts of the Archives & Special Collections (ASC) unit at the University of Glasgow to preserve the large, at-risk collection of the NVA Archive. We discuss the nature of the collection, and the way it was used to evaluate our digital preservation capability.

Keywords – Digital preservation, at-risk collections, cultural heritage, University archives

Conference Topics – Community

I. INTRODUCTION

The Archives & Special Collections (ASC) unit at the University of Glasgow Library is responsible for "managing, promoting, enabling access and supporting engagement with the Library's unique and distinctive collections"[1]. ASC collects and provides access to archival records, manuscripts, rare books, and other primary and secondary sources to support teaching and research at the University and the wider community. The University has been an early adopter of digital records and processes [1] – and following an ongoing digital transformation as part of the University's "World-Changing Glasgow Transformation" initiative[2] ASC has been increasingly collecting born-digital records of historical, cultural and business significance.

This paper will discuss our efforts to build a robust digital preservation service through the lens of our work to preserve one digital collection with continuing value to the Scottish arts community (and beyond), the NVA Archive. NVA - an acronym for 'nacionale vita activa' (roughly translated as 'the right to influence public affairs') [2] – was an internationally renowned arts organisation based in Glasgow, which closed in 2018. NVA epitomized community participation through public art that "reconnects people to their built and natural environment" taking inspiration from the "Ancient Greek ideal of a lively democracy, where actions and words shared among equals bring new thinking into the world." [3]

The size of the NVA digital records, and the impetus to secure their ongoing existence after the closure of the organisation, pushed our (rather nascent) digital preservation processes and systems to their limits; but also helped us evaluate our capability to preserve large, at-risk digital collections.

II. ABOUT THE NVA ARCHIVE

NVA was founded in 1992 by Angus Farquhar, and has produced a number of key public art projects including *The Hidden Gardens* in Glasgow, *The Storr* on the Isle of Skye, the international touring work *Speed of Light*, and the *Kilmahew/St Peter's* project based around the rescue of the modernist ruin of St Peter's Seminary[3].

In September 2017, the Directors of NVA (Europe) Limited made the decision to withdraw from plans to rescue St Peter's Seminary. Over the following months, attempts were made to develop an alternative proposal for the building and to stabilize the organisation, but it became clear that this was not possible. Additionally, in February 2018, NVA

[1] www.gla.ac.uk/myglasgow/archivespecialcollections/
[2] www.gla.ac.uk/myglasgow/worldchangingglasgow/
[3] Information on NVA artworks available at: http://nva.org.uk/artworks/

18th International Conference on Digital Preservation
iPres 2022, Glasgow, Scotland.

received notification that its bid to Creative Scotland for core revenue funding had been unsuccessful. The scale of the financial challenges had become untenable and, in June 2018, the Board of NVA announced that it was to close.

ASC approached the Business Archives Surveying Officer at the Ballast Trust - a charitable foundation that provides a rescue, sorting and cataloguing service for business archives[4] - for assistance in securing temporary storage and a records survey for the records, as the collection was deemed at-risk due to the NVA's imminent closure. In August 2018, the physical records along with a hard drive of digital records were transferred to the Ballast Trust, where they were box-listed by the Surveying Officer.

The digital records amount to approximately 800GB, of which 300GB are images. The remaining records consist primarily of word documents, excel spreadsheets, PDFs, and images in various file formats (see Table 1). The files themselves include correspondence, minutes, financial records, administrative records, staff records, images, video, audio, project plans, project evaluation, tender and funding bid documents, marketing, and promotional materials.

Table 1 Volume and composition of the NVA digital archive

Data volume	*782GB*	
File formats	40,000 image files 11,579 PDF 19,501 documents 5,820 spreadsheets 248 presentations	150 audio files 359 video files 878 Adobe files 117 AutoCAD files 337 zip files

III. Community and Archival Value

NVA has been at the heart of the Scottish contemporary art world for over two decades. The organisation had a track record of producing large scale public performance artworks which thousands of people in Scotland (and internationally) have engaged with. Their last, unfinished project, Kilmahew/St Peter's, involves Scotland's best known modernist building and one of the first 'modernist ruins'. This, along with the enduring legacy of many of their past works, makes it highly likely that the records of NVA will be of significant interest to art and architectural historians, geographers and the art and cultural sector. Recognizing this value, our primary aim has been to preserve all that we can of the NVA archive. This approach will provide an opportunity for a community of voices to inform decision-making around appraisal, description, and access. It has, however, also challenged our digital preservation capacity across staffing, resource, and technology.

IV. Digital Preservation in ASC

To enable long-term preservation and continuing access to digital records, the University of Glasgow Library has been investing into digital preservation via the Digital Preservation Working Group (DPWG)[5], a cross-University collaboration working to implement the University's digital policy and strategy. Established in 2015, the group oversees the delivery of digital preservation services, with representation from the University Library, IT Services and the Data Protection & Freedom of Information Office. The DPWG is also responsible for setting, maintaining and monitoring compliance with the University's Digital Preservation Policy [4].

For Archives & Special Collections, the drive to develop our digital preservation service is multi-faceted. There is, however, a strong connection to the communities we are part of. The University community – both as an organization and a community of researchers and students is key. However, we also sit within the wider heritage, cultural, and business communities. As a collecting institution, it is critical that we engage creatively and practically with the digital to ensure that we can continue to document, preserve and provide access to Scotland's economic, cultural, and creative heritage.

Within the business heritage community, ASC has worked for many years in partnership with the Business Archives Surveying Officer and the Ballast Trust to facilitate the survey, rescue, and preservation of business archives. Increasingly, as in the case of the NVA Archive, this work is centred around digital records. Reflecting on our work to preserve the NVA Archive, we will highlight the impact this Archive has had upon various aspects of our capacity to support the preservation of born-digital business records.

V. Evaluating Digital Preservation Capability

The digital records of the NVA Archive provided an excellent case to evaluate the capability of our

[4] https://ballasttrust.org.uk/

[5] https://bit.ly/3Cms6gN

18th International Conference on Digital Preservation
iPres 2022, Glasgow, Scotland.

digital preservation system, a cloud-hosted instance of Archivematica 1.12.2 with 1TB cloud storage for processing and archival. In particular, we were keen to explore the system's performance with collections of high volume, file numbers and format diversity.

We used our digital archiving workflow[6] (Figure 1) as the basis for evaluation. The workflow formalises and amalgamates existing ASC processes and practices around collections development, acquisition and appraisal, with digital preservation processing requirements.

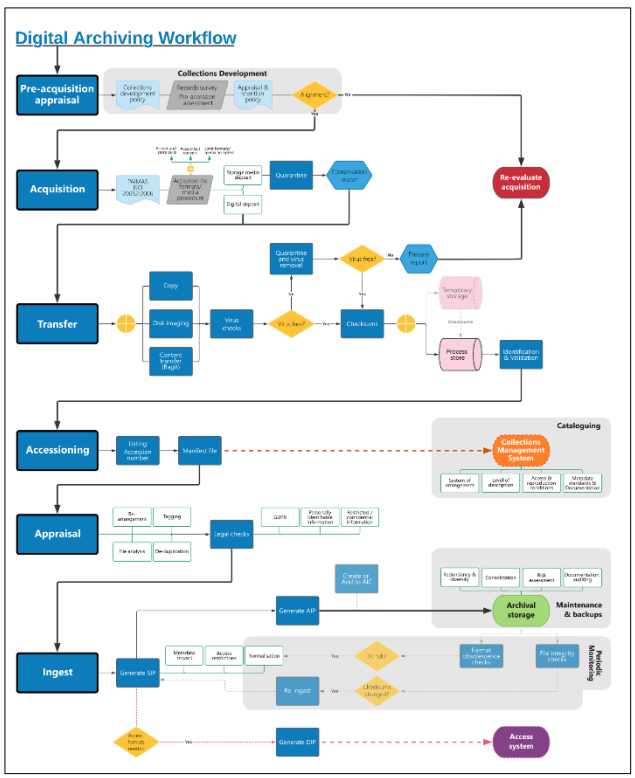

Figure 7 The ASC digital archiving workflow

PRE-ACQUISITION APPRAISAL

As a Glasgow-based limited company by guarantee, NVA fits the university's Collection Policy which states that "Archive Services primarily seeks to acquire records of Scottish business in the 19th, 20th and 21st centuries." [5]. The records survey conducted by the Ballast Trust, further highlighted a number of Intellectual Property and Data Protection issues which informed legal checks and the content of rights metadata.

EVALUATION METHODOLOGY

We identified twenty test cases across six areas relating to preservation processing: cloud storage performance; system administration; transfer; appraisal; ingest; and archival storage. The latter five

Figure 8 Matrix of impact and likely for risk assessment

Likelihood \ Impact	Negligible	Minor	Moderate	Significant	Severe
Very likely	Low	Medium	High	High	High
Likely	Low	Medium	Medium	High	High
Possible	Low	Low	Medium	Medium	High
Unlikely	Low	Low	Medium	Medium	Medium
Very unlikely	Low	Low	Low	Medium	Medium

areas were meant to match the respective Archivematica tabs. Table 2 summarises the tests that were conducted per area.

Table 2 Summary of tests conducted per preservation processing area

Cloud storage performance	Test different methods of uploading files onto cloud storage
	Administer files on cloud storage
Transfer (Archivematica)	Complete automated and interactive transfers using default and custom processing configurations
Appraisal (Archivematica)	Use appraisal tools for content analysis and create SIPs from appraised/re-arranged content
Ingest (Archivematica)	Prepare and store AIPs for SIPs using preservation planning strategies for normalization
Archival storage (Archivematica)	Manage AIPs in the Archival storage tab
Administration (Archivematica)	Manage requests for AIP deletion and recovery
	Add, edit and delete Processing configurations
	View and manage failure reports
	View and manage Processing storage usage

Issues identified during testing were assigned risk levels, which were calculated using matrix of **Impact** and **Likelihood** of an identified Issue occurring and disrupting digital preservation processing (Figure 2).

[6] The latest version of the workflow is available on the Community Owned Workflows (COW) of COPTR: https://bit.ly/3MsArUG

18th International Conference on Digital Preservation
iPres 2022, Glasgow, Scotland.

The level of risk was assessed by the potential of an Issue to trigger a threat event, which in turn could give rise to one or more of the following:

Loss of Functionality

- The system does not perform one or more of its intended functions, either partly or entirely.
- The system exhibits unexpected behaviour(s) which affect or inhibit completion of operations.

Loss of Integrity

- The system and its data can no longer be trusted.
- The systems and its data are incomplete or incorrect.
- The security and confidentiality of the system and its data have been compromised (e.g. unauthorised access, wrong user permissions).

Loss of Availability

- The system can no longer be accessed.
- The system does not respond to valid queries and/or produces system fault errors.

The risk assessment was used to classify issues as either Low, Medium or High Risk (see Table 3).

Table 3 Risk assessment - levels of risk

Low	Medium	High
The Issue pertains to a non-critical or supplementary service/function/operation.	The Issue pertains to a core service/function/operation.	The Issue pertains to a critical and/or exigent service/function/operation.
The Issue does not prevent completion of operations.	The Issue may prevent completion of operations, but there are workarounds.	The Issue prevents completion of operations.
The Issue affects functions/operations that are easily recoverable or reproducible.	The Issue affects functions/operations, for which alternatives exist.	The Issue affects functions/operations that cannot be recovered or reproduced; and the related effects cannot be otherwise mitigated.
The Issue does not cause loss of integrity or availability.	The Issue does not cause loss of integrity but affects aspects of availability.	The Issue can cause loss of integrity and/or availability.

Evaluation Results

A total of 19 issues were documented across the six preservation processing areas. The risk assessment conducted showed that 45% of high-risk issues related to cloud storage performance, followed by issues of Administration (22%). Similarly, 29% of all issues related to cloud storage performance; and 24% with appraisal (Table 4).

Table 4 Summary statistics of issues per risk level and preservation processing area.

	High	Medium	Low	% of total
Cloud storage	44%	0%	33%	29%
Transfer	11%	20%	0%	12%
Appraisal	11%	20%	67%	24%
Archival storage	11%	40%	0%	18%
Administration	22%	20%	0%	18%

With minor exceptions, the majority of the issues encountered during testing derived from the large volume and number of files in the NVA archive. We tested uploading files from the acquired hard drive onto cloud storage via both a web interface and a desktop client. For uploads via the web interface, server timeouts occurred in all tests where the file size was larger than 200MB. Uploads via the desktop client had a higher success rate, but sync speed was slow: in one test, it took 2 days to upload ca. 200GB of data. In other cases, file integrity checks failed due to file corruption during upload.

Archivematica performed more consistently across tests but struggled when transfer volume exceeded 1GB or when a transfer consisted of more than 1,000 files.

VI. Lessons Learnt

The experience of preserving the NVA Archive highlighted a number of areas where our digital preservation service needs improvement. Although the digital archiving workflow – and related ASC processes and policies – provided a good foundation for digital preservation processing, technical issues with cloud storage impeded operations.

Workarounds exist to mitigate the risks identified during testing. For instance, the number of files per transfer can be limited, and the collection can be preserved as multiple AIPs. However, problems persist with large files – e.g. a single video file in the NVA Archive was almost 3GB.

As we continue to develop our digital preservation service, and the more we engage with the digital preservation community, we expect to further our understanding of these issues and find robust solutions to preserve our at-risk collections.

18th International Conference on Digital Preservation
iPres 2022, Glasgow, Scotland.

ACKNOWLEDGMENT

As highlighted above, Archives & Special Collection's work to preserve the NVA Archive has been a partnership between the University and the Ballast Trust's Business Archives Surveying Officer. Chris Cassells, now of the National Library of Scotland, undertook a records survey of the NVA in 2018, and subsequently completed initial digital preservation actions for the collection across 2019 and 2020. His professional, and practical, input to the management of the Archive is greatly appreciated.

REFERENCES

[1] E. Yan and C. Paterson, "World-changing Transformation: Collections at University of Glasgow Archives & Special Collections", 15th International Digital Curation Conference (IDCC20), 2020. https://bit.ly/3MxZztj

[2] Fairweather, "DIANE WATTERS, St Peter's, Cardross: Birth, Death and Renewal", *Architectural Heritage*, vol. 27, no. 1, pp. 119-122, October 2017.

[3] NVA, "Annual report 2015/2016", February 2017. https://bit.ly/3pOiNkC

[4] University of Glasgow, "Digital Preservation Policy", v3.8, 2020. https://bit.ly/3CDpMlP

[5] University of Glasgow, "Collections Development Policy", v1.0, February 2017. https://www.gla.ac.uk/media/Media_591723_smxx.pdf

18th International Conference on Digital Preservation
iPres 2022, Glasgow, Scotland.