iPRES 2025


iPRES 2025
WELLINGTON
AOTEAROA
NEW ZEALAND
BnF
cfibd
comité français international bibliothèques et documentation
Kia Ora !
Bonjour !
Hi !
iPRES 2025 Wellington
Yannick.Grandcolas@bnf.fr
Expert Preservation Numérique @ Bibliothèque nationale de France
& Directeur @ Open Preservation Foundation

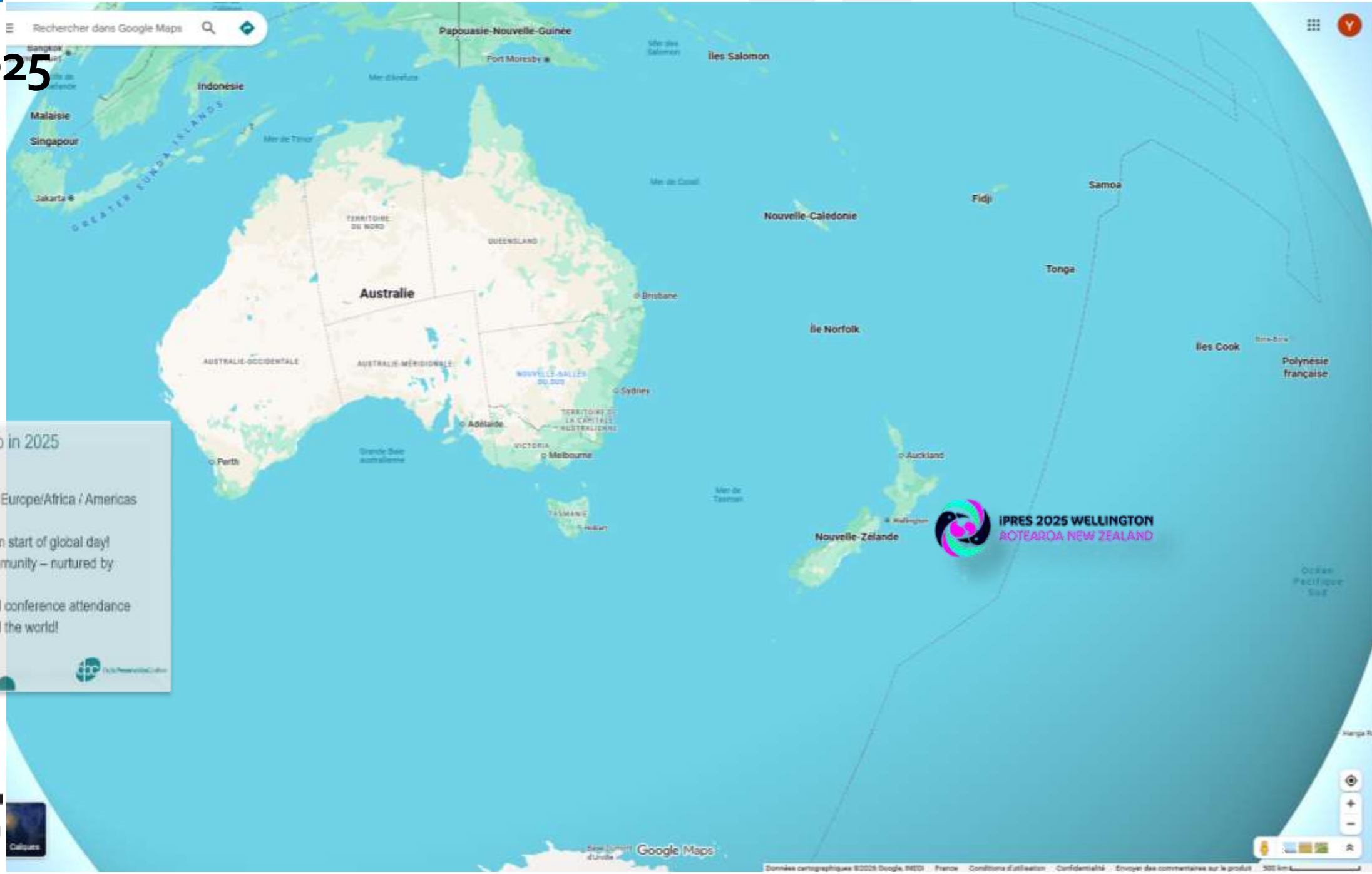
iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

- +600 participants
- +50 pays représentés
- +250 contributions, articles, présentations, posters, tables rondes, ateliers, tutoriels, réunion des Bibliothèques nationales…

https://zenodo.org/communities/ipres2025recordings/records?q=&l=list&p=1&s=10&sort=newest

BnF

iPRES 2025



iPRES Radio in 2025

- Now key part of iPRES
- 3 x global sessions – Australasia / Europe/Africa / Americas
- Multiple hosts
- Robin/Josh running iPRES TV from start of global day!
- Seed planted in global iPRES community – nurtured by conversation/collaboration
- A new way of thinking about virtual conference attendance
- Equitable participation – all around the world!

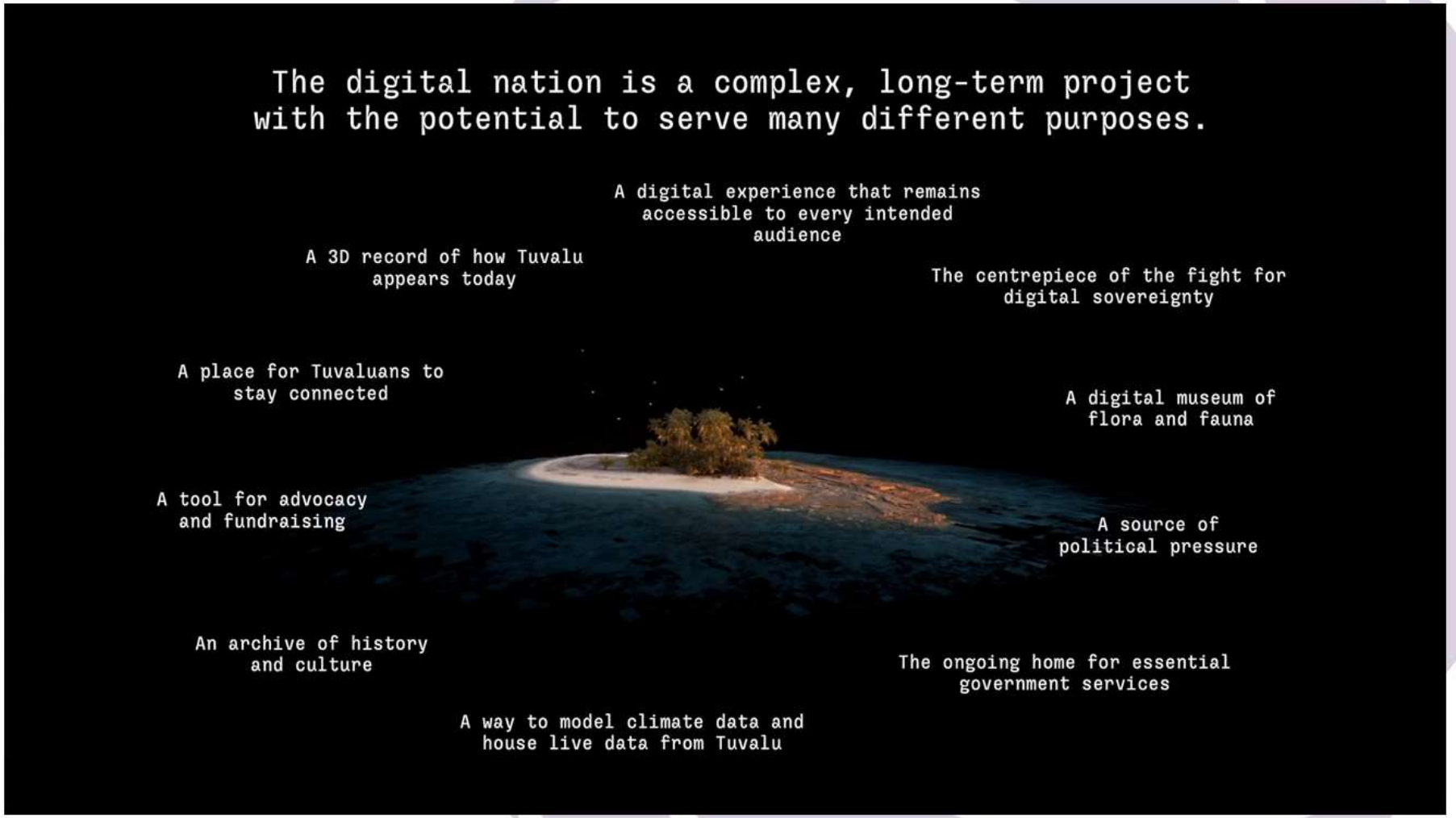
{BnF

iPRES 2025

La Nation numérique

Construire l’architecture du futur de Tuvalu

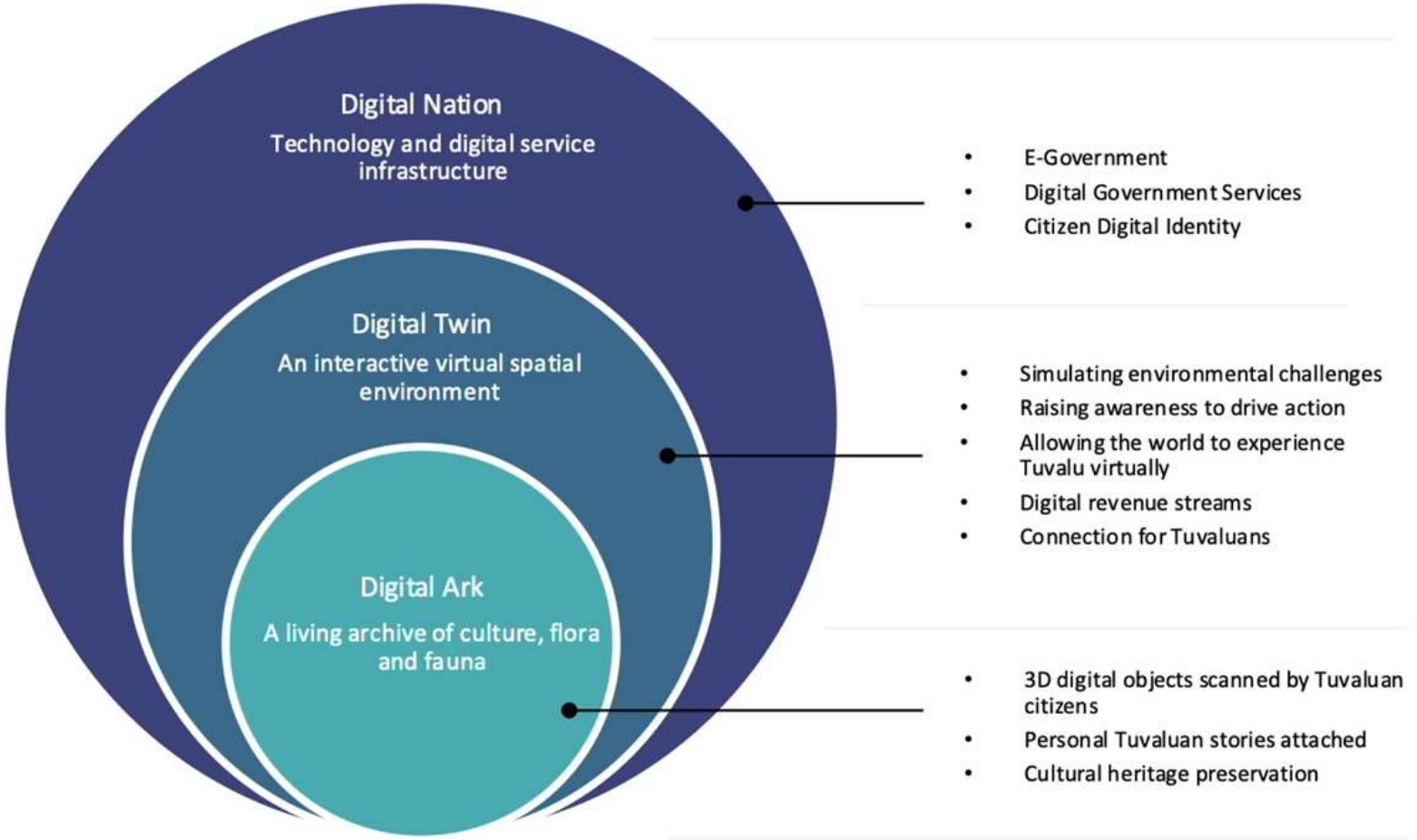
Simon Kofe


The digital nation is a complex, long-term project with the potential to serve many different purposes.
A digital experience that remains accessible to every intended audience
A 3D record of how Tuvalu appears today
The centrepiece of the fight for digital sovereignty
A place for Tuvaluans to stay connected
A digital museum of flora and fauna
A tool for advocacy and fundraising
A source of political pressure
An archive of history and culture
The ongoing home for essential government services
A way to model climate data and house live data from Tuvalu


iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND
21ST INTERNATIONAL


THE FIRST
DIGITAL
NATION


iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND
21ST INTERNATIONAL

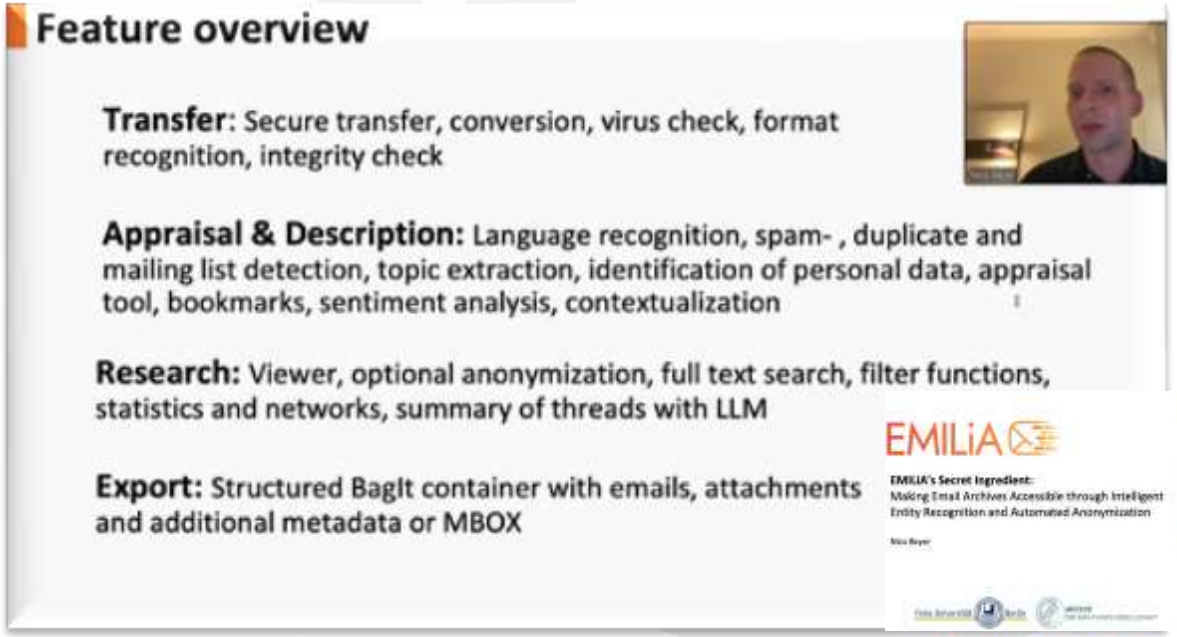

Digital Nation
Technology and digital service infrastructure
Digital Twin
An interactive virtual spatial environment
Digital Ark
A living archive of culture, flora and fauna
E-Government
Digital Government Services
Citizen Digital Identity
Simulating environmental challenges
Raising awareness to drive action
Allowing the world to experience Tuvalu virtually
Digital revenue streams
Connection for Tuvaluans
3D digital objects scanned by Tuvaluan citizens
Personal Tuvaluan stories attached
Cultural heritage preservation

iPRES 2025

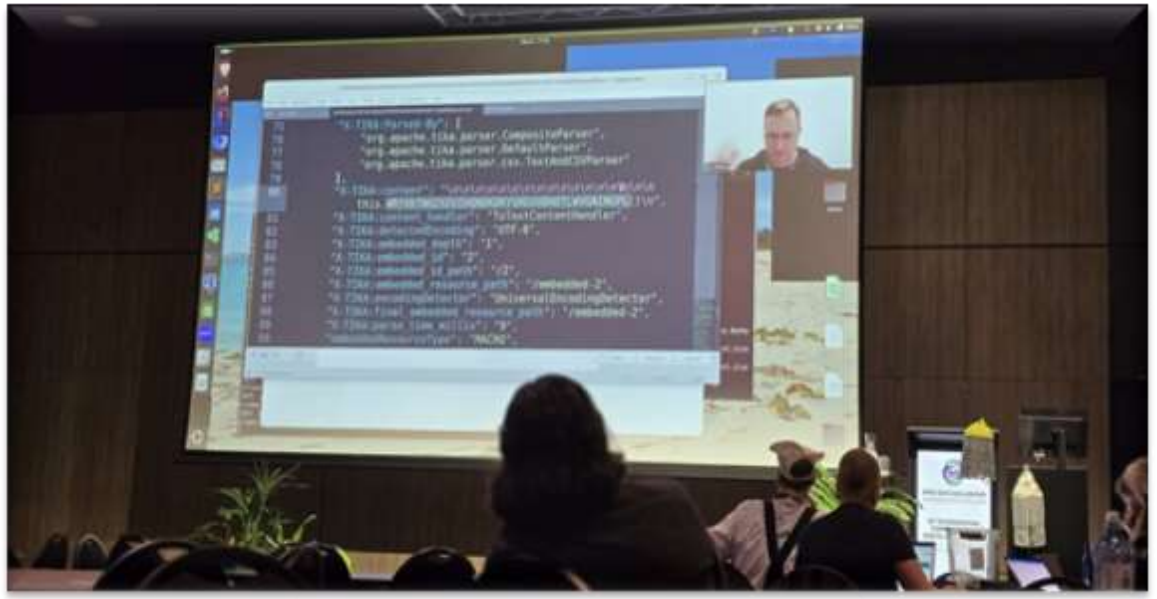
iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Bibliothèque nationale Wellington
60e anniversaire 1965-2025

The Game of LOCKSS
Stanford U.
The Game of LOCKSS | LOCKSS Program
Danielle Taylor Indiana U.



Emilia's Secret Ingredient : Préservation d'Archives courriels grâce à l'IA et un traitement semi-automatisé (anonymiser…) Nico Beyer – U.Berlin.



A taste of Apache – Tika. Tim Allison, expert – Président & VP

BnF

iPRES 2025

Contributions BnF

Table ronde :

BnF

iPRES 2025

Autres contributions France

FRiPRES 2025 est la seconde édition de la journée francophone dédiée à la préservation numérique, organisée en parallèle de la conférence internationale iPRES 2025 (Wellington, Nouvelle-Zélande).

Cette journée offre aux **professionnels, chercheurs et étudiants** un espace pour

Morgän Attias, Infotel

How To Persist a Persistent Identifier? : Long Paper - iPRES 2025 - Wellington, New Zealand / Te Whanganui-a-Tara, Aotearoa

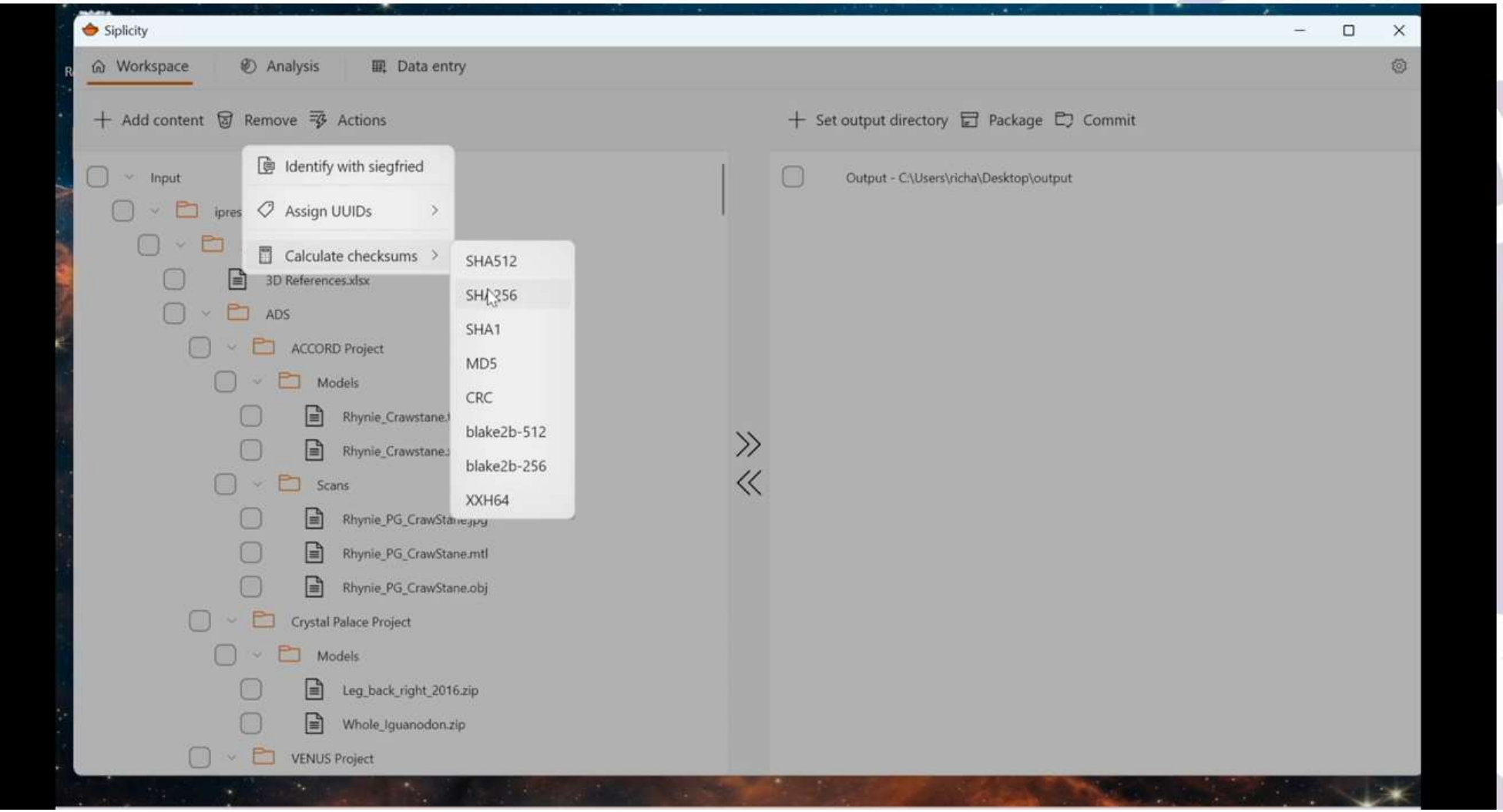
19.02.2026

Marion Ville

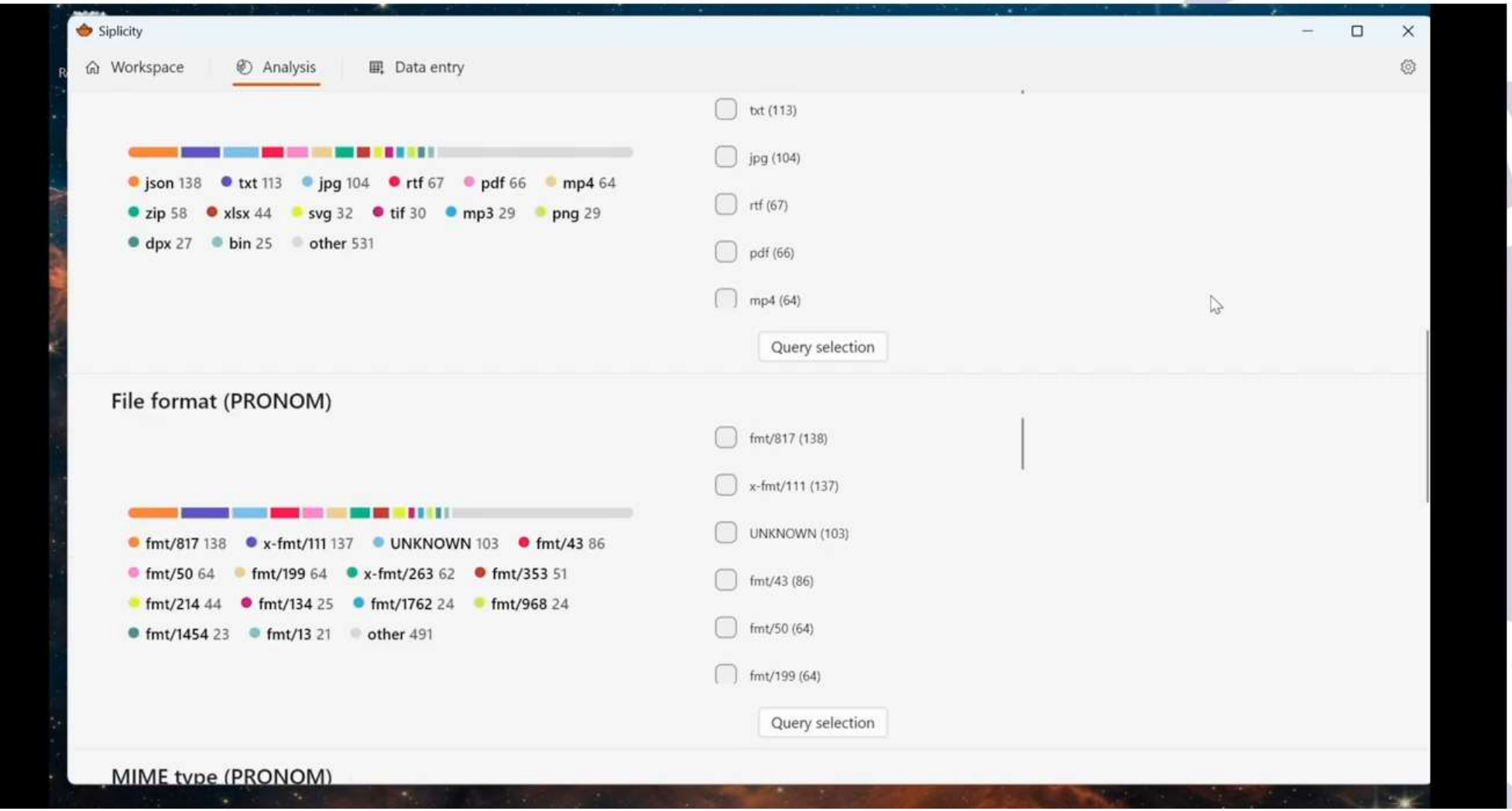
Many French archival services use persistent identifiers to facilitate access to archives. However, few studies have addressed the sustainability of these identifiers. This can raise questions when these identifiers are employed to identify digital archives preserved in digital repositories.
The Vitam program, which develops the Vitam digital recordkeeping software on behalf of the French government, carried out a study in collaboration with the Vitam user community in 2021. The goal was to improve knowledge of the use of persistent identifiers, particularly ARK identifiers, as well as their lifecycle and the requirements regarding their sustainability. The study also aimed to define the features required in the Vitam software to guarantee a long-term preservation of this type of identifier, whatever happens. Finally, it also made sure that their use could be adapted to different user contexts.

https://phaidra.univie.ac.at/o:2198818

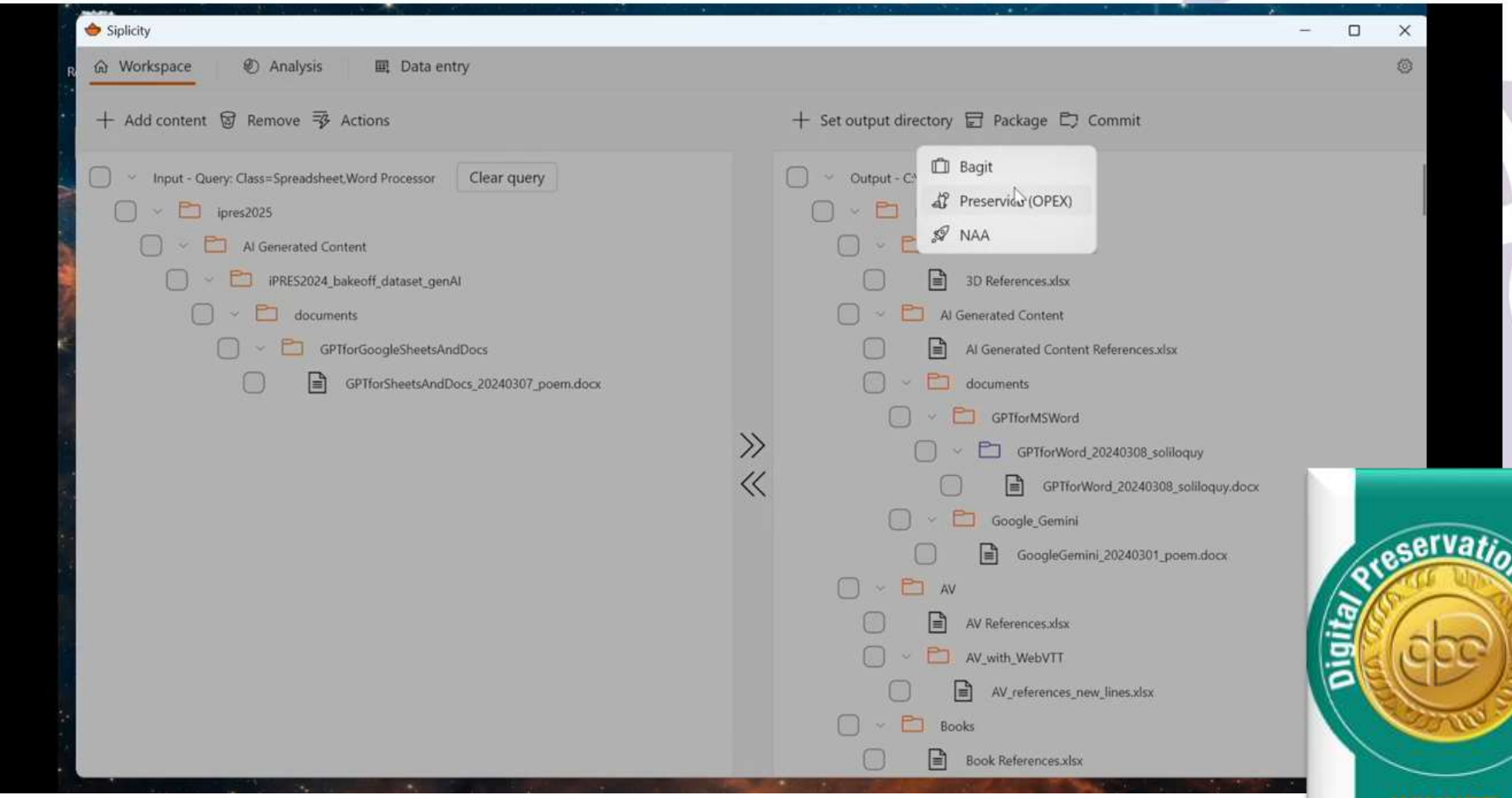
BnF



Siplicity (Fair Core License)
Richard Lehane, Consultant Préservation Numérique, Sydney, Australie.



Siplicity (Fair Core License) Richard Lehane, Consultant Préservation Numérique, Sydney, Australie.



Siplicity (Fair Core License)
Richard Lehane, Consultant Préservation Numérique, Sydney, Australie.

iPRES 2025

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Published November 2025 | Version v1

Conference paper | Open

Unlocking Web Histories: Leveraging LLMs and RAG to Transform Discovery in Web Archives

Davis, Corey[1]

Show affiliations

This paper explores how Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) can be applied to improve access to web archives. These collections are often difficult to navigate due to their complexity and the limitations of traditional search tools. The author examines how RAG can help address concerns around trust and transparency in AI by grounding LLM outputs in external, curated sources. At the University of Victoria Libraries, a custom RAG pipeline was developed to build on tools like WARC-GPT. This pipeline enables natural language querying with source attribution and incorporates optimizations in data preprocessing, chunking strategies, and hardware acceleration. When tested on real-world web archives of cultural significance, it demonstrated improved retrieval accuracy and computational efficiency. The discussion also considers the broader implications for digital preservation in an age of uncertainty, emphasizing the importance of trust, ethics, sustainability, and continued human oversight in AI-powered discovery. The findings suggest that RAG offers a promising way to unlock the value of underused digital heritage collections while upholding the foundational values of research libraries, including long-term preservation, equitable access, and responsible stewardship of historic materials.

Series information

Full Paper

Files

iPRES2025_LP_Unlocking_Web_Histories-Davis.pdf

1 of 11 | Automatic Zoom | Download

1 of 11

UNLOCKING WEB HISTORIES: LEVERAGING LLMS AND RAG TO TRANSFORM DISCOVERY IN WEB ARCHIVES

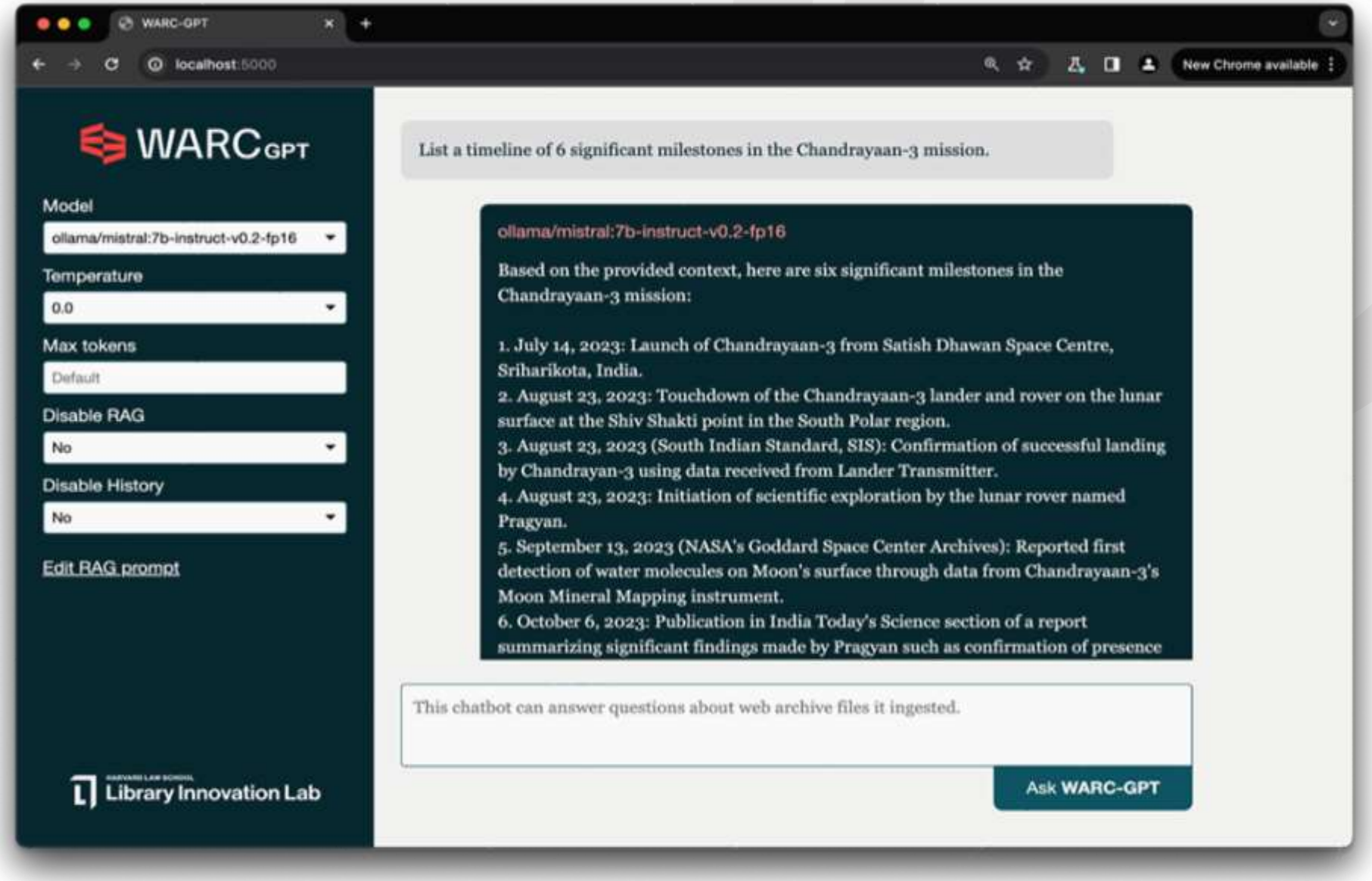
Corey Davis
University of Victoria
Canada
coreyd@uvic.ca
https://orcid.org/0000-0002-6542-5865

Article : Unlocking Web Histories: Leveraging LLMs and RAG to Transform Discovery in Web Archives
Davis, Corey - Victoria U. - Canada.

BnF

iPRES 2025

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND



Article : Unlocking Web Histories: Leveraging LLMs and RAG [Retrieval-Augmented Generation] to Transform Discovery in Web Archives
Davis, Corey - Victoria U. - Canada.

BnF

iPRES 2025

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

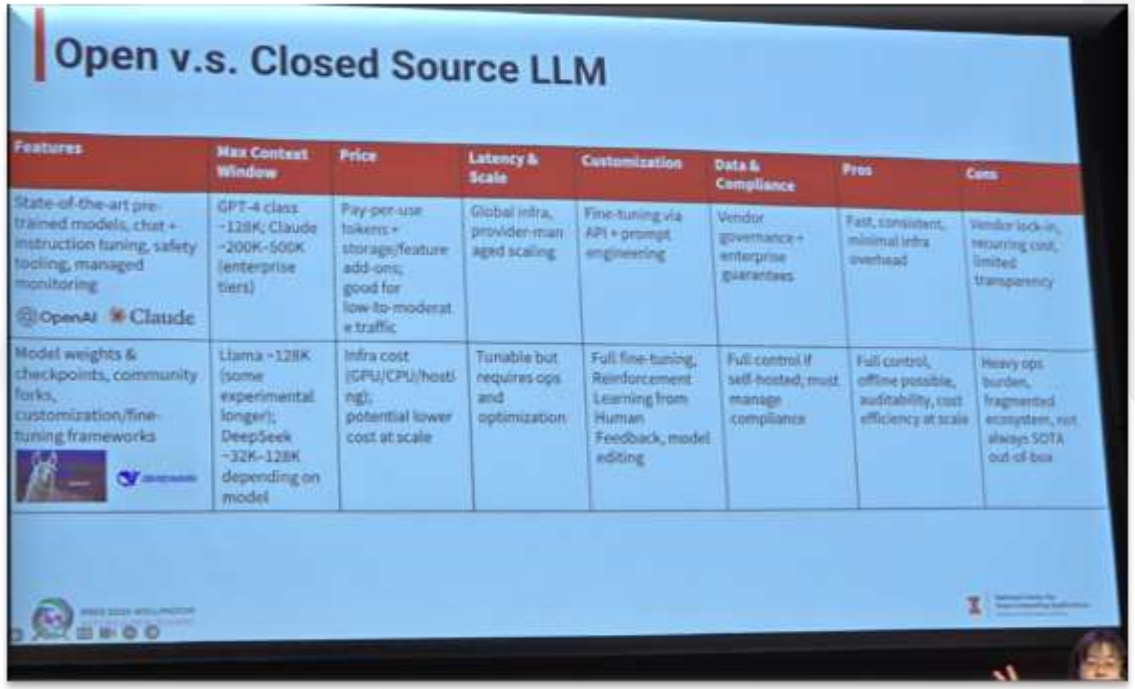
Atelier : Intégration d'un Grand Modèle de Langage Web dans les systèmes de recherche
Integrating LLM Web Functionality in Search Systems : Minimized Coding to Create a Smart Search Web Interface
Pengyin Shan, Ingénieure de Recherche Logiciels (Illinois U.)

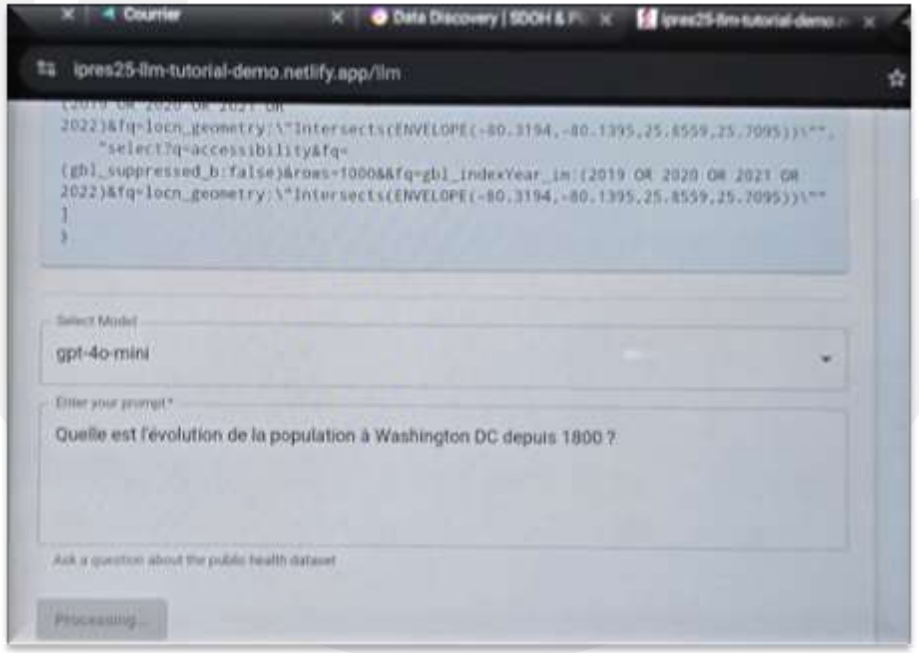
Description : Build Smart Search Interfaces: Use Prompts to Turn Questions into Solr Queries with Minimal Coding

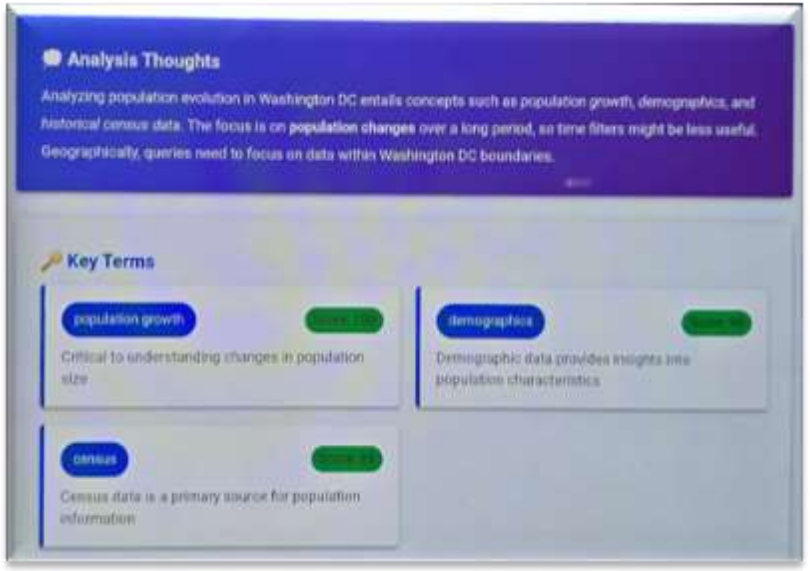
{BnF

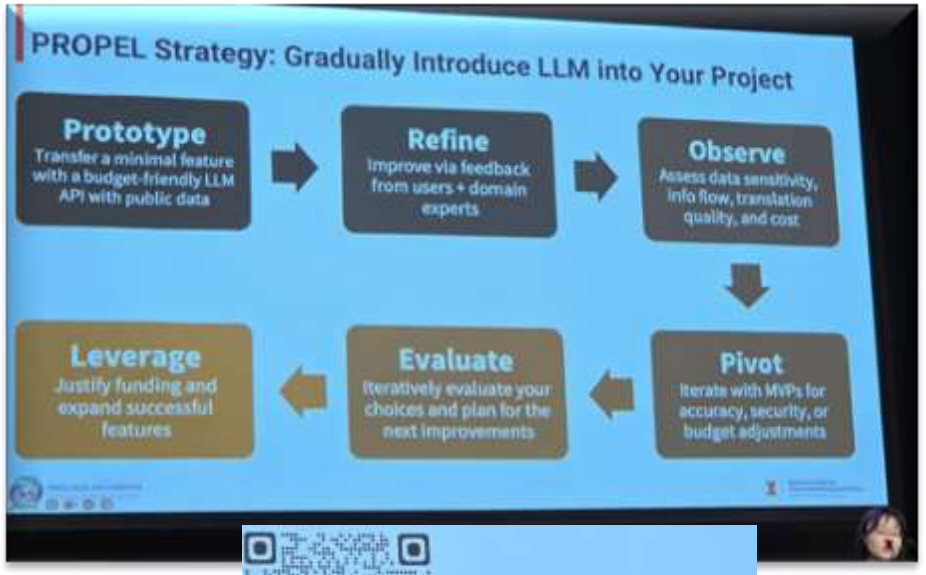
iPRES 2025

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND









{BnF

Atelier : Intégration d'un Grand Modèle de Langage Web dans les systèmes de recherche
Integrating LLM Web Functionality in Search Systems : Minimized Coding to Create a Smart Search Web Interface
Pengyin Shan, Ingénieure de Recherche Logiciels (Illinois U.) Contact :

shorturl.at/8A7H8

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

iPRES 2025

Atelier ARK : identifiants persistants

John Kunze
https://arks.org/about/ark-faq-fr/
& California Digital Library

Archival Resource Key (ARK) persistent identifiers: affordable long-term citation and access in the age of GenAI

ARK Alliance
arks.org

John Kunze, Donny Winston
ARK Alliance, 2025

What are ARKs used for?

- genealogical records (12 billion FamilySearch)
- publisher content (157 million Portico)
- scientific datasets and records (22 million INIST)
- scanned books and texts (50 million Internet Archive)
- bibliographic records (27 million BnF main catalog)
- museum specimens (15 million Smithsonian Institution)
- public health documents (20 million UCSF IDL)
- historical documents (36 million CDL, 6 million BnF Gallica)
- historical authors and scholars (4 million SNAC)
- fine art museum collections (490,000 Louvre)
- vocabulary terms (30,000 Periodo, YAMZ)

Why care about ARK identifiers?

- Because robust web links are rare – the average URL lifetime is 100 days
- ARKs can be "persistent identifiers" (PIDs), which serve as permalinks
- "Ten persistent myths about persistent identifiers"
  https://n2t.net/ark:13030/c7gb1xh09

The ARK (Archival Resource Key) identifier scheme was introduced in 2001.

French National Library ARKs: 12148

https://gallica.bnf.fr/iiif/ark:12148/btv1b8449691v/**f29/2131,4016,1467,948/full/0/default.jpg**

region coordinates
full quality
page
file format

Tools

Documentation and Software:
arks.org/resources

- Minters and resolvers: Noid, arknoid, arklet, and **arklet-Frick**
- Other minters: counters, UUID, ULID, …
- Journal minter/resolver: OJS Plugin
- In-house library ARK system: **ARKs Service UTScarborough**
- Consider Suffix Passthrough in the style of arklet-Frick and EZID

Test création ark : https://ezid.cdlib.org/

Dictionnaire communautaire avec identifiants persistants :
https://yamz.net/

BnF

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

iPRES 2025

Atelier : Calcul de l'empreinte carbone de la préservation Numérique

An Approach to Estimating the IT-Hardware Carbon Footprint

- Hardware (servers, other storage devices) consist of components that have an in-body carbon footprint reported by vendors
- PAIA (Product Attribute Impact Algorithm) is a source for all hardware carbon footprint reports
- A consortium of hardware vendors develops and reports carbon footprint statistics (e.g. Cisco, IBM, Hewlett Packard Enterprises, Dell and Lenovo)
- These values can be turned into a generic server components in-body carbon footprint calculation for components with missing vendor figures

The Carbon Footprint Calculation Formula

7.5 PiB platform, 3+1 copies

$CFP_{tot} = \Sigma (Q + E \cdot PUE \cdot C)$

Our CFP: 21169 kg CO_2ekv

- CFP_{tot} - Total annual carbon footprint
- Q - Carbon footprint of hardware manufacturing, disposal and logistics (per year)
- E - Electricity used to run the DPS platform
- PUE – Data centre power usage efficiency
- C - Electricity production carbon footprint

Our Data Centres: A: 1.02 B: 1,2 C: 2

Finnish hydropower 24 kg CO_2ekv/ MWh

The Basis For Calculating a Generic Carbon Footprint Model

- More & more hardware component vendors report their carbon footprint statistics for their hardware components, e.g. Seagate
- A generic estimate for all server or tape library hardware components can developed by taking median value from existing reported values
- In-Body = Hardware manufacturing, logistics and disposal
  - Enclosure 15 kg CO2ekv
  - Fan 3 kg CO2ekv
  - Power supply unit 19 kg CO2ekv
  - Mainboard (2x CPU + 192GB RAM) 488 kg CO2ekv
  - Solid state drive (<500GB) 19 kg CO2ekv
  - Hard drive 37 kg CO2ekv
  - Daughter-board 170 kg CO2ekv
  - LTO9 tape drive 217 kg CO2ekv
  - LTO9 tape media 7.5 kg CO2ekv
  - IBM TS4300 21U tape library 6035 kg CO2ekv

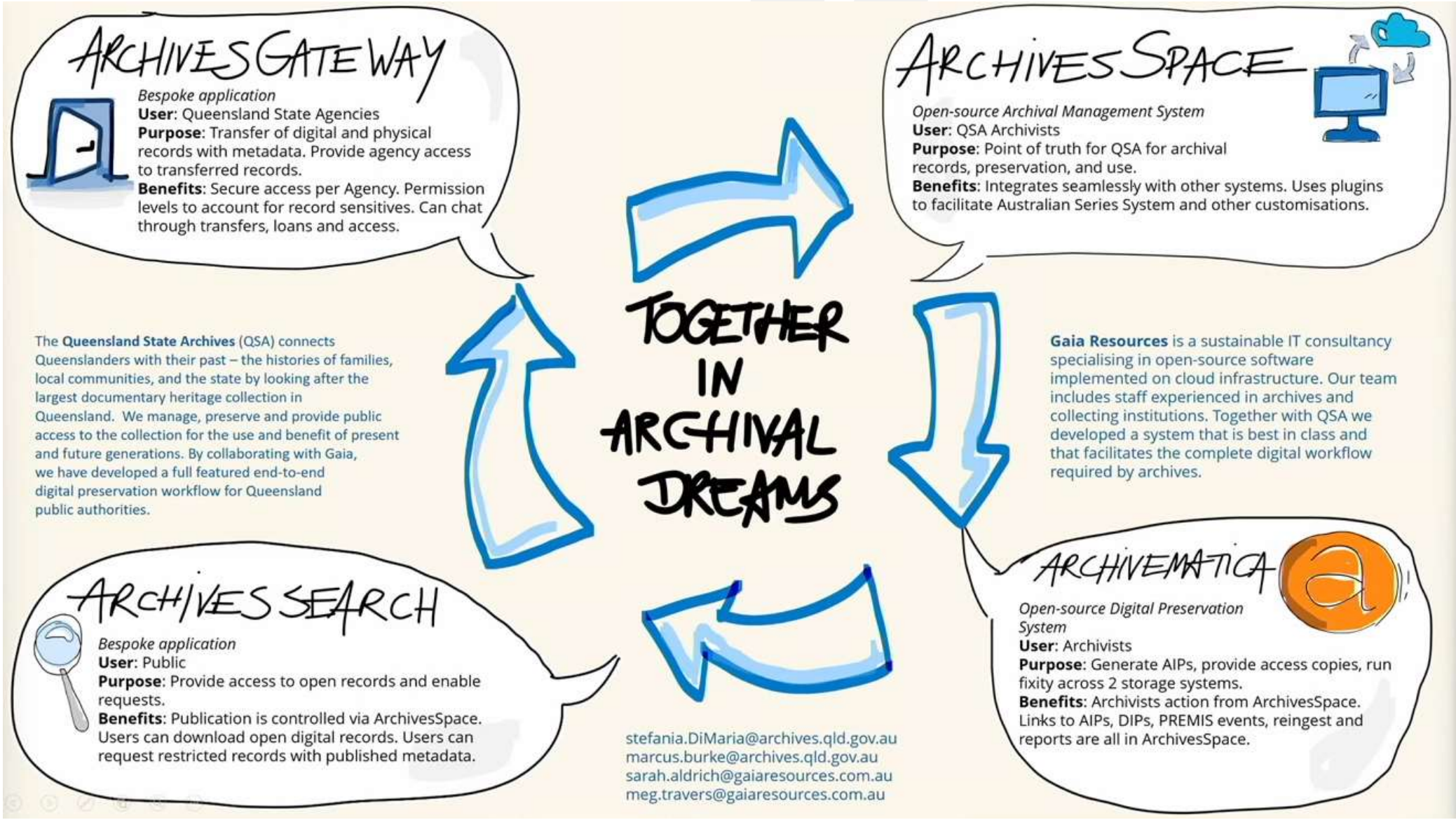
- Feedback from the exercises
  - What figures are missing?
  - Other environments/setups that should be considered?
  - Other aspects to consider?
- What next?
  - Work on CFP calculation continues within the carbon footprint task force led by the Digital Preservation Coalition (DPC): https://www.dpconline.org/
  - The DPS in Finland has been a member of DPC since 2021
- Please contact us if you wish to discuss carbon footprint calculation!

BnF

iPRES 2025 Posters

Stefania di Maria & Marcus Burke (Archives Queensland, Australie)

Sarah Aldrich & Meg Travers (Gaia Consultants Open Source)



iPRES 2025 Posters

iPRES 2025 WELLINGTON AOTEAROA NEW ZEALAND

iPRES 3 to 7 November 2025, Wellington New Zealand

BUILDING A SINGAPORE DIGITAL PRESERVATION COMMUNITY OF INTEREST TO FURTHER ADVOCACY AND ENGAGEMENT

NATIONAL library SINGAPORE
Charlyn Ang, Librarian
charlyn_ang@nlb.gov.sg

First a Practitioner, Then a Leader

The National Library Singapore's digital preservation policy sets out our commitment to digitisation and digital preservation of Singapore's documentary heritage and **long-term strategic goal of being a leader in the field of digital preservation.**

In that vein, the National Library provides staff with opportunities to further **develop their skills with courses, seminars, conferences and self-directed learning**, ensuring all relevant **staff are appropriately skilled to carry out digital preservation activities.**

DPC's Robin Wright speaking on 15 May 2024

NLB joined the Digital Preservation Coalition (DPC) in 2021 to actively engage with the wider community and learn from other professional bodies, and has been **advocating digital preservation practices by engaging practitioners and the general public.**

Between 2021 and 2024, the National Library has organised **in-house talks** as well as actively **collaborated with DPC** experts to hold a workshop and talks, sharing knowledge on digital preservation topics and amassing a total of over 1,000 participants over the 4 years.

Deepening Collaboration with the New Singapore Community of Interest

From left: Speakers from NUS Libraries, Heritage Conservation Centre and National Library Board moderated by National Library Singapore"

To bring engagement to the next level, the National Library Singapore spearheaded a new **digital preservation community of interest** amongst local tertiary libraries and heritage institutions with the objectives to:

1. **Build a thriving local community that:**
   - **Supports, encourages**, and **engages** all members.
   - Inspires **active participation** in digital preservation efforts.
   - Encourages members to **reinvest their time in advancing preservation** within their own organisations.
2. **Facilitate sharing between practitioners from different disciplines, enabling members to:**
   - Deepen collaboration through **networking.**
   - Gain access to **shared experiences, resources, and tools.**

The inaugural session hosted a panel of speakers from **National University of Singapore Libraries, Heritage Conservation Centre and the National Library Board Singapore** who presented "The Digitisation Journey & Experience of NUS Libraries", "Preserving Time-Based Media Art in the National Collection", and "Digitisation and Digital Preservation Workflow" in the respective order.

It was followed by a **moderated Q&A** session that:

1. Fostered **multilateral engagement between members of the audience and the panel.**
2. Provided a **multidisciplinary exchange of insights and perspectives.**

Inaugural Session: Reception and Feedback

69 participants from a mix of government, heritage, and tertiary institutions participated in the session.

What did you find useful in this session?

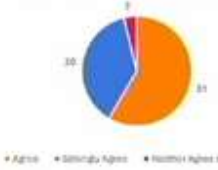
"Hearing about how digital preservation is done for different fields, especially for time-based media and research data."

"Frank look at the difficulties of digital preservation and how considerations can cut across the different presenters."

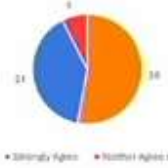
"A stimulating discussion on the merits and limitations of cloud storage in the digital preservation space, as well as deeper insights into the specificities of time-based media preservation."

"The sharing by personnels from different organisations broaden my perspective on practices adopted by different organisations."

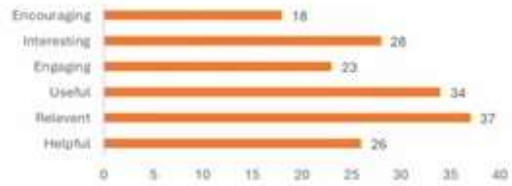
This session encouraged me to further develop my understanding and practice of digital preservation.



The session provided clarity and deepened my understanding of digital preservation.



I found this session:



The response was positive with clear interest indicated for other topics, showing **potential for future sessions.**

The success of the session can be attributed to the **open communication** between the speakers and audience through the moderated Q&A, encouraging candid sharing of **information outside of the panel's presentations** and opportunities to get **immediate answers from experts** present.

Impact survey

A follow-up survey was conducted in September to investigate the networking and learning impact of the session.

Respondents were encouraged to build connections with speakers and other participants, **30%** shared they were able to network with others during or after the event, and **50%** responded they intend to form connections in future sessions.

Average scores out of 5

Results show a healthy interest in topics such as digital preservation workflows, formats, metadata and systems.

80% like sessions with **moderated Q&A**, and **27%** have indicated interest in **being a speaker** for upcoming sessions.

With these results, the foreseeable longer-term impact of the community of interest include:

- **More collaboration** by engaging with peers from other institutions.
- **Improved trust** with increased transparency in practices, fostering a **culture of openness** where institutions feel safer to share methods, challenges, and ideas.
- **Lowered barriers to learning** with access to valuable expertise and insights, **raising the overall skill level and maturity of digital preservation efforts** in Singapore.

BnF

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Discover your Guinness Story

Repurposing a digital preservation system into a historical database

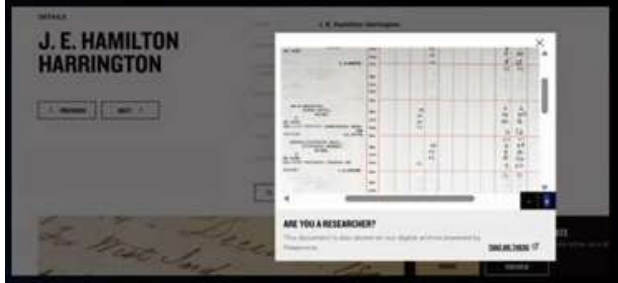
The Guinness Archive partnered with Ancestry in 2023 to digitise and transcribe 1,315 personnel and trade ledgers. The project resulted in 250,000 digital images (7TB), and over 1.6 million individual data entries. The Archive embarked on a project to create a user-friendly database within a digital preservation environment. Digital preservation systems represent a significant investment for any memory institution and this provided the Archive with an opportunity to use it to meet additional business needs.

Considerations

Preservica only system at Archive's disposal that could comfortably store and process 7TB.

Internal stakeholders request that access was via the Guinness Storehouse website. Maximum load time of website needed to be maintained.

Create user-friendly genealogical database, returning transcribed information and digitised images.



GUINNESS ARCHIVE

Approach

Ingest
Package digital objects and metadata using Preservica's Open Preservation Exchange (OPEX) format and Preservation Asset eXchange (PAX) protocol for ingest. Link to Axiell Collections catalogue.

Create metadata fragments
Create a metadata fragment for each line of data and package it with its associated digital object (or ledger image). The nature of the ledgers as day-to-day records meant that a single image could be associated with tens of individual fragments.

Develop API
Fragments made the images searchable within Preservica using the personal names, addresses and dates. An API call was developed using Preservica's developer tools to run searches entered by the user in the GSH website interface.

Stumbling blocks
Initial results from this approach were mixed. The API was successful in being able to retrieve metadata fragments but, in cases where a fragment field was empty, data would be 'bumped up' from the next fragment, creating mismatches.

A different approach - Data as record
Re-ingest each line of csv data as an individual 'record' within Preservica and link to digital image via UUID. We mapped the API search to the fields we wanted to be publicly searchable, and testing proved that this solution had been successful.

Image retrieval
Initial attempts to display images using third-party image viewers were unsuccessful. The web tech team were eventually able to use the API to generate a thumbnail representation of the image hosted in Preservica.

Test & Launch
Search went live in March 2025 following several rounds of internal testing.

Results

Significant success in terms of harnessing available resources to create a user-friendly historical database for our users.

Increased accessibility of collections.

Bolsters the business case for continued investment in digital preservation system. Diversifies functionality.

Average monthly visitors: 1,112, avg. 1m and 14s
Peak monthly visitors: 15,017, avg. 1m and 52s

Learnings

Delivery of data
Define preferred delivery method of project outputs from the outset.

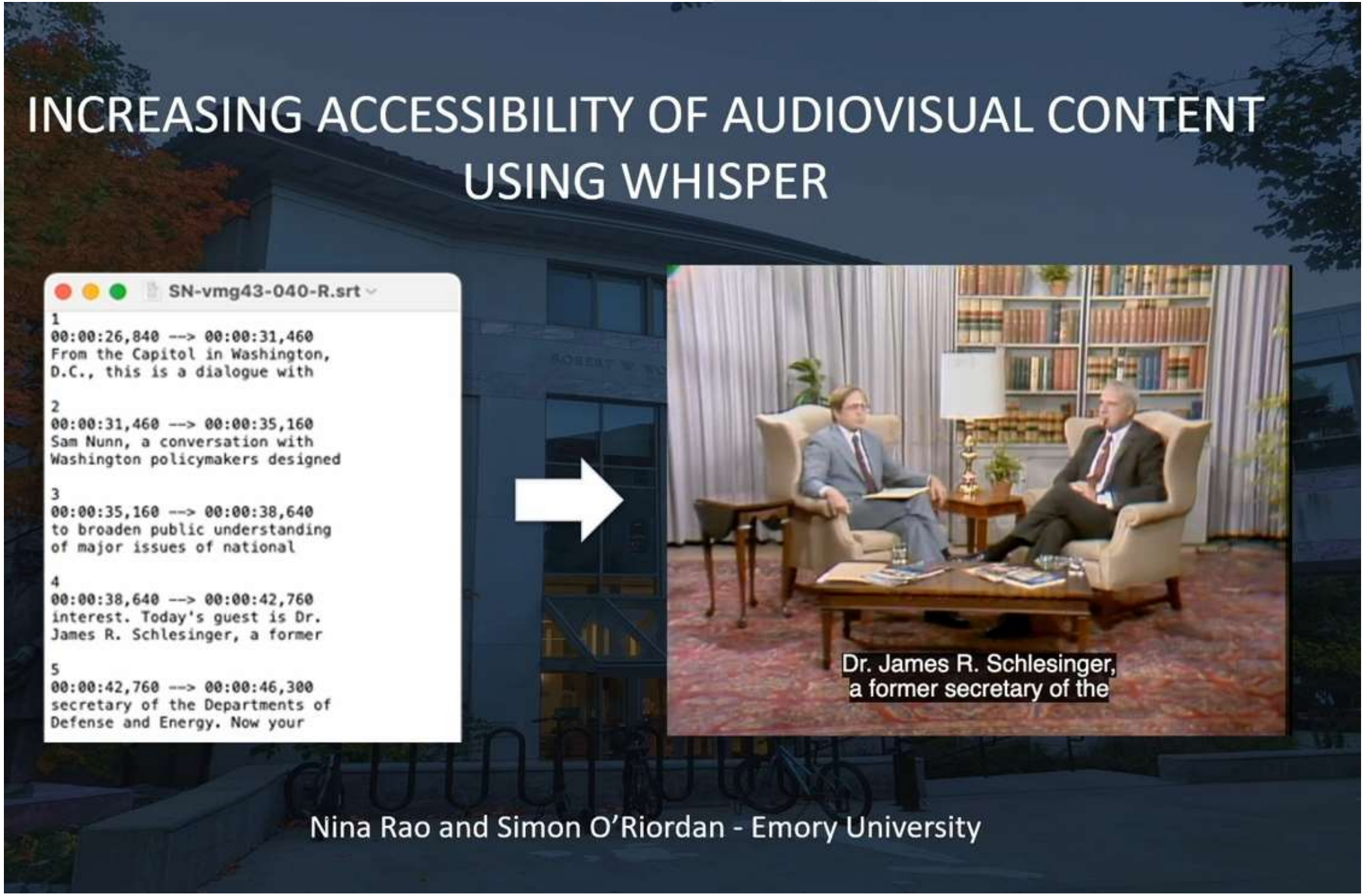
Third-party communication
Ensure that any developers remain in constant communication with the vendor. Digital preservation software is bespoke and specific functionalities or quirks may not be immediately apparent to IT professionals. A lot of trial and error was avoidable.

Leanne Harrington, Digital Archivist
leanne.harrington@diageo.com

BnF

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND



https://openai.com/fr-FR/index/whisper/

BnF

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Open Preservation Foundation

ONE YEAR OF "BY JHOVE! EXPLAIN YOURSELF!"

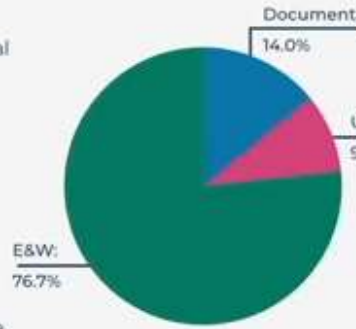
JHOVE is one of the most widely used tools in digital preservation, but anyone who uses it will be familiar with the various difficulties that follow. Causes of these problems vary, ranging from not understanding the impact of validation errors to outdated training materials to a graphical user interface (GUI) considered not user-friendly. "By JHOVE! Explain Yourself!" is a Special Interest Group (SIG) hosted by the Open Preservation Foundation (OPF) that brings together members of the OPF as well as interested parties within the digital preservation community who are interested in evolving JHOVE and discussing their issues with peers.

FIRST YEAR RESULTS

12 Monthly virtual meetings

4 Quarterly presentations

43 People on the mailinglist



Entries on Task Spreadsheet

18 Average number of regular attendees

28 Number of institutions on the mailinglist

86 Total entries spreadsheet

THE THREE PILLARS

Documentation

The JHOVE documentation was outdated and challenging for new users. This pillar focuses on revising existing content and developing new guidance materials.

The team began by converting the documentation from Java to Markdown to improve integration and collaboration. Since then, internal guides and live GitHub tutorials tailored to JHOVE have been created.

Documentation updated:

- An Introduction to JHOVE
- Getting Started with JHOVE
- PDF-HUL module
- JPEG-HUL module

Errors and Warnings (E&W)

Users found JHOVE's error & warning messages, especially for complex formats like PDF, overwhelming and unclear. This pillar addresses specific issues. The team have worked on analysis, walk-throughs, and updated the JHOVE GitHub Wiki with their findings.

Some of the updated E&W:

- TIFF-HUL-3: Unknown data type
- TIFF-HUL-6: Tag count mismatch
- TIFF-HU-26: StripOffsets not defined
- PDF-HUL-45: Malformed filter
- PDF-HUL-85: No document catalogue dictionary

User-friendliness

Using JHOVE requires patience and support; the web application, desktop application, and command line are all slightly different, and the results are rarely cohesive. Recognising this, members of the group have developed a new GUI prototype, which we hope to implement in the next year.


GUI Prototype

WHAT'S NEXT

- Catering to other timezones
- Providing better documentation
- Visualise the progress
- Further develop GUI
- Make a roadmap
- Join us!

Micky Lindlar
micky.lindlar@tib.eu
TIB Leibniz Information Centre for Science and Technology

Lotte Wijsman
lotte.wijsman@nationaalarchief.nl
National Archives of the Netherlands

Georgia Moppett
georgia.moppett@openpreservation.org
Open Preservation Foundation

iPRES 2025

BnF

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

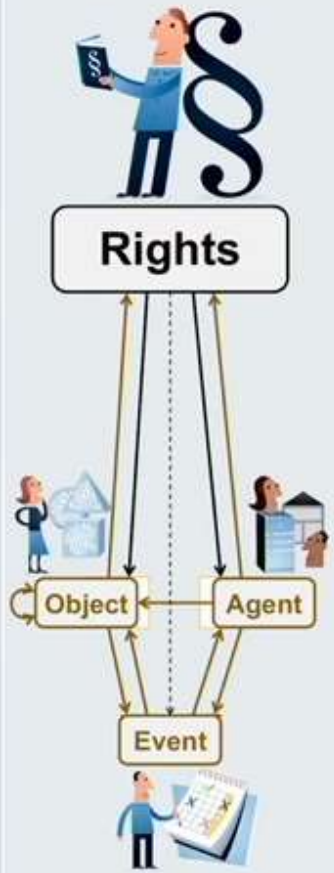
BnF

PREMIS
PREservation Metadata Implementation Strategies

New RIGHTS in PREMIS

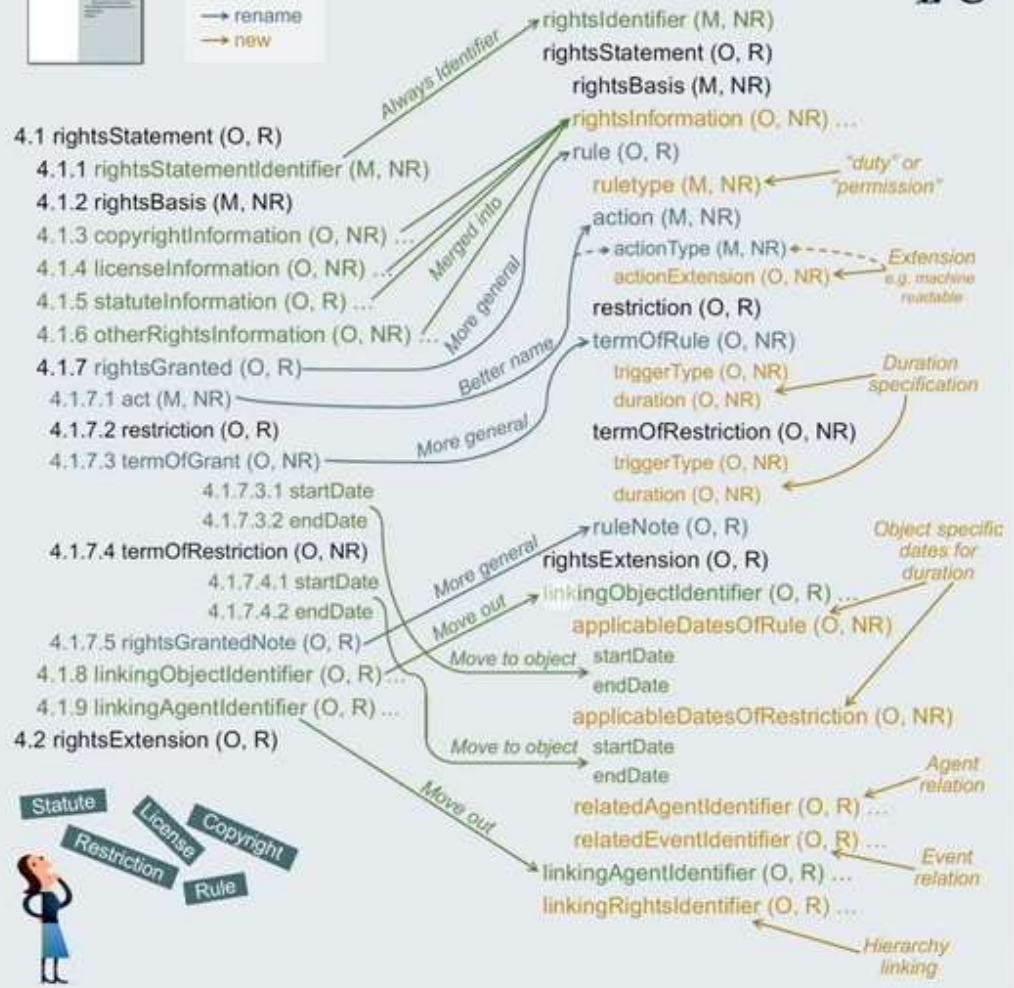
Enhanced options to manage RIGHTS metadata

More about PREMIS:
https://www.loc.gov/standards/premis/

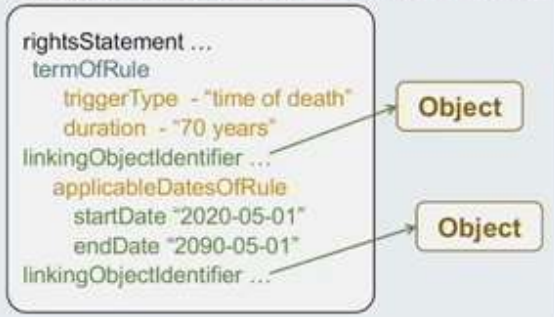


Changes in Data Model

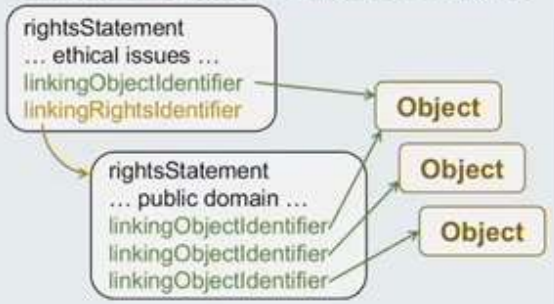
Legend
- move
- rename
- new



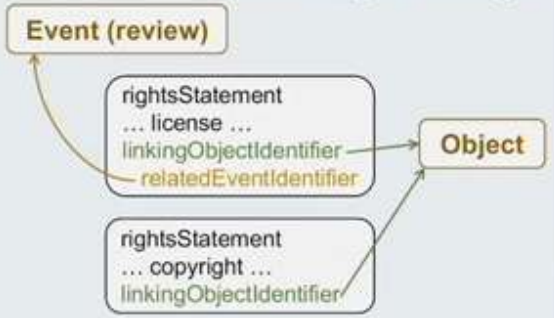
Embargo for objects, for duration of years



Different Levels of Conditional Rights



Review event leads to rights change



Marjolein Steeman
Netherlands Institute for Sound & Vision

Leslie Johnston
National Archives and Records Administration

Karin Bredenberg
Kommunalförbundet Sydarkivera

Micky Lindlar
TIB Leibniz Information Centre for Science and Technology

Eld Zierau
Royal Danish Library

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

{BnF

King Abdullah University of Science and Technology
جامعة الملك عبدالله للعلوم والتقنية

Using Claude 3.7 Sonnet + Project to Build A University Archives Metadata Management System

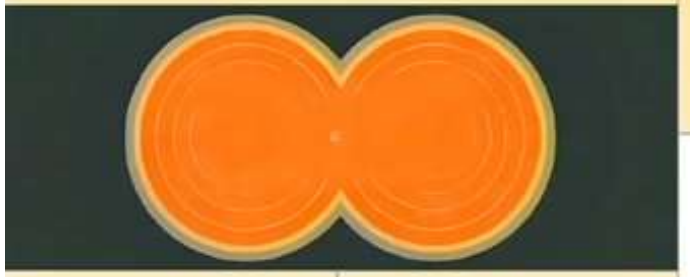
Request. Implement. Troubleshoot. Repeat.

Eamon Smallwood
University Archivist, King Abdullah University of Science and Technology
Saudi Arabia
Eamon.smallwood@kaust.edu.sa
https://orcid.org/0000-0001-9557-5816

Acknowledgements
Claude 3.7 Sonnet was used to develop documentation and code for the KAUST University Archives Metadata Management System.
Poster images created with Canva (https://canva.com).

Inspiration

This project was inspired by the work of Hudhayfa Nazoordeen, a 20-year-old student at the University of Waterloo who built a proof-of-concept nuclear fusor in his bedroom. Nazoordeen's post describing the project stated that he had "zero hardware experience," his secret of success was "Claude Sonnet 3.5 + Project." Our project challenge, an order of magnitude less ambitious, is to build a proof-of-concept, functional, and integrated University Archives Metadata Management System (UAMMS). As AI tools increase in capability, it is valuable to investigate how far a traditional archivist with limited development experience can go in building a comprehensive metadata management service. Additionally, it has been a valuable learning experience in the strengths and weaknesses of the current generation of AI tools.



AI-Assisted Development Strategy

- Claude + Project
  - Key feature - project documentation repository - 200,000-word capacity
- Preconditioning
  - Novice-level explanations
  - Powershell syntax
  - Proof of concept project - LIMIT FUNCTIONALITY

Project Requirements

- Support the Dublin Core (DC) metadata schema
- Support the creation of new metadata records
- Support updating existing metadata records
- Support import and export using Excel and XML

Implementation

The UAMMS sandbox achieved remarkable development speed, progressing from initial concept to a working database with a web interface in under 5 hours of active development on the first day. Key insights from the development and implementation process:

- **Rapid initial development**
- **Metadata normalization challenges**
- **Steep API integration learning curve**
- **Enhanced deployment speed**
- **Persistent realignment of AI code is required**

Technical Architecture

- **Front-end: HTML, CSS, JavaScript with Bootstrap for responsive design**
- **Backend: Python Flask web application**
- **Database: SQLite with SQL Alchemy ORM**
- **API: RESTful API for integrations**

We focused on a simple approach and a standard set of prerequisites to support scaling up the UAMMS once the proof of concept build was stable. Claude generated the code base directory as the hub for all application and user interface functionality. The code is publicly available on GitHub. (https://github.com/eamonsmallwood/UAMMS)..

Prompt Samples

Results and Next Steps

Positive:

- **Rapid code development and deployment**
- **Quick remediation of simple code errors**
- **Code development crash course**

Negative:

- **AI code frequently beyond scope of the request**
- **Low continuity of AI code**
- **AI code is challenging to maintain and update**

Next Steps:

- **Stable deployment of UAMMS**
- **Continued integration and fine-tuning**

Conclusion

Claude 3.7 Sonnet + Project proved to be an effective method for supporting the development of a metadata management tool with minimal programming experience required. This project demonstrates that the current state of AI-assisted code development is more beneficial for those with limited programming experience. Experienced programmers may find the code generated by current LLMs overly elaborate and requiring many revisions. Claude tended to be unnecessarily verbose when developing product documentation and code. Claude required a persistent stream of limiting prompts to refine the outputs to fit the project scope and requirements. Even as LLMs and other AI tools continue to advance, a human actor must be involved in specifying requirements and evaluating whether the AI outputs meet those requirements.

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Digital Preservation Storage Criteria Update

Informing a New Usage Guide for the Community

History

- **2015:** Working group formed
- Workshops, feedback, presentations
- **2016:**
- **2017:** 2nd version of criteria published
- **2018:**
- **2019:** 3rd version of criteria + usage guide published
- **2020-2021:**
- 4th version of criteria

Digital Preservation Storage Criteria

Criteria	Category	Description	Related Standards
Integrity checking	Content integrity	Performs verifiable and/or auditable checks to detect changes or loss in or across copies (e.g. checksum recalculation, fixity checking, identifying missing files)	ISO 15801 ISO 16363
Cost-efficient	Cost considera-tions	Costs relatively less overall than other comparable solutions, by being designed with cost efficiencies, for example, has resource pooling and sharing, multi-tenancy (multiple users share the same applications)	ISO 16363 ISO 17797
...	...	...	...

Standards mapping

Standard	Description
ISO 14721	Reference model for an archival information system (OAIS)
ISO 16363	Audit and certification of trustworthy digital repositories
...	...

Play the Game

Learn about the Digital Preservation Storage Criteria using the game.

Learn more

Read more about the Digital Preservation Storage Criteria, and participate in the development

Usage guide

Motivation: Using the Digital Preservation Storage Criteria requires context that is grounded in digital preservation principles

Existing topics:

Risk management
Mitigate loss of data through risk management

Elements in establishing bit safety
Number of copies, Independence between copies and Integrity checks

Independence between copies
Mitigates risk that an event, an agent or technology can harm several copies of data

Costs
The more effectively risks are mitigated, the more preservation storage may cost

Potential New topics

Documentation and Logging
This section includes the following topics:
- Change management processes
- Documentation of infrastructure components and inter-relationships
- Audit trail documentation
- Logging data documentation
- Business continuity and disaster recovery documentation

Preservation Strategy and Technology Review
This section covers the following topics:
- Technology migration
- Format migration
- Media obsolescence
- Energy usage management and planning
- Technology Watch/Environmental Scans

Organizational Operations
This section includes the following topics:
- Organizational planning processes
- Service level agreements (SLAs) and Contracting processes
- Organizational policies for management of data
- Organizational policies for management of infrastructure
- Organizational roles and responsibilities
- Organizational resilience planning processes

System Security and Auditing
This section covers the following topics:
- Preservation storage security management
- Integration of security for preservation storage with other relevant systems
- Auditing and audit trail requirements

Storage Media Management
This section covers the following topics:
- Media traceability and supply chain documentation
- Physical media selection and management
- Media storage conditions

Vocabulary/Glossary
Commonly used terms and their sources

Give you vote and comment!

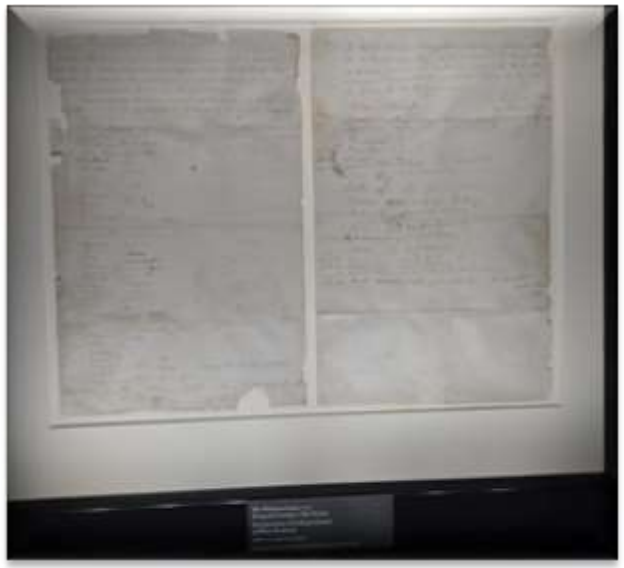
Eld Zierau, Royal Danish Library
Cynthia Wu, National Library of New Zealand
Nancy Y McGovern, Global Archivist, LLC
Sibyl Schaefer, University of California

BnF

iPRES 2025

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Visites professionnelles
Bibliothèque nationale

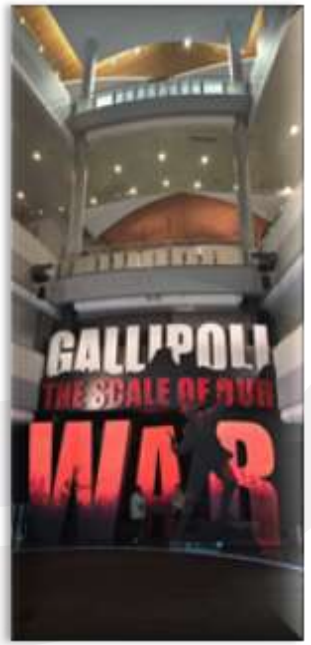


{BnF

iPRES 2025

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Musée national
& Parlement



BnF

iPRES 2025

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Merci !

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Building the Data-driven Archive

Scaling-up and automation with ETL and asynchronous event-driven messaging architecture

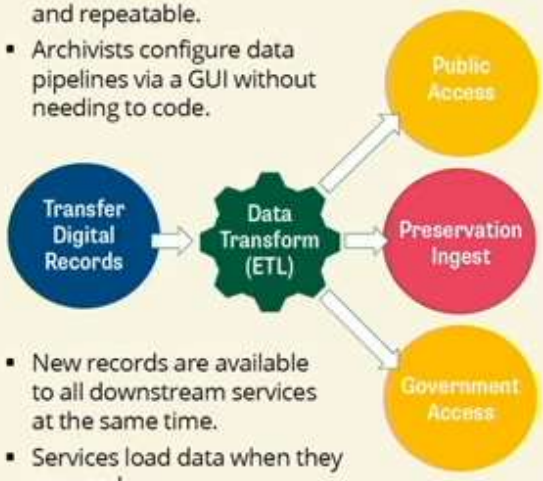
THE NATIONAL ARCHIVES

Problem: Linear process, every service transforms data

Proposal: Centralised, automated, data infrastructure

First: Non-linear data flow

- Data tranformations are centralised, audited, and repeatable.
- Archivists configure data pipelines via a GUI without needing to code.



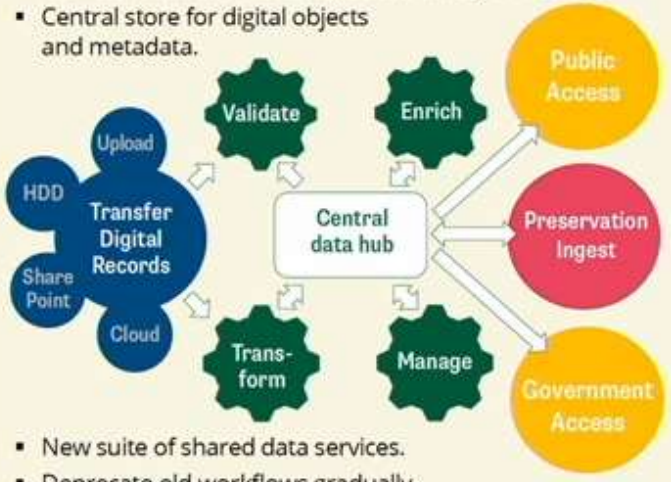
- New records are available to all downstream services at the same time.
- Services load data when they are ready.

Now: Event-driven orchestration

- Message flows initiate work and signal task completion.
- Activities are discrete and asynchronous.
- Message database enables reporting.

Next: Central data, new data services

- Add workflows to handle a greater variety of records.
- Central store for digital objects and metadata.



- New suite of shared data services.
- Deprecate old workflows gradually.

Further iteration: Data catalogue for machine-readable insight into content & flows

Future direction: Data governance for re-usable business rules & logic

Phil.Ashworth, Sonia.Ranade, Paul.Young
@nationalarchives.gov.uk

BnF

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Risk Register Revival: Preservation Planning in a Time of Change

The management of a digital archive requires the ability to audit and interrogate digital formats for a variety of archival functions. Some examples where format analysis work can beuseful are Transfers advice, digitisation format selection and format migration for preservation or access purposes . The National Archives of Australia uses the 'performance model' as a guideline for digital archiving decisions that means the archive strikes to preserve the meaningful content of archival records rather than only focusing on preserving the original format. In 2022 The National Archives developed a File format register and risk matrix for digital file formats as part of a broader strategy to wholisticaly manage the digital collection. This register was created to centralise management of a dispersed and varied collection that is held across multiple systems. The risk matrix has been evaluated and updated by a newly formed team and the file format register and reporting process is undergoing current uplift made possible by system and data connectivity improvements.

Background: A Risk Framework for Digital Formats

The National Archives accepts and seeks to actively and openly manage risk in all its activities.

In 2022 the Preservation section developed A Risk Framework for Digital records to support both organisational capability and preservation planning. The Framework was designed to work with the National Archives Risk Management Guide to ensure alignment with risk approaches at an organisational level.
It also incorporated aspects of the ABC method of Risk Assessment developed by the International Centre for the Study of the Preservation and Restoration of Cultural Property (ICCROM). and the risk management cycle described by ISO 31000:2009 Standard for Risk Management. The approach sought to incorporating high level collection risk assessment using tools the Digital Preservation Coaliton RAM (Rapid Assessment Model) as well as developing an-inhouse risk assessment matrix to enage with assessing formats in the collection to inform preservation planning.

Review: Identified areas for improvement

Format Register

The original format register was created to manage and prioritise preservaton for a digital collection that is described and held across multiple systems and locations. It has a section to include both internal and external format risk assessments and identification methods including the PRONOM register and NARA's format evaluations and planning. The triage section was taking too much staff time and required streamlining and areas of the register used for reporting were now able to be improved in light of a large scale Archives-wide project improving and unifying collection data.

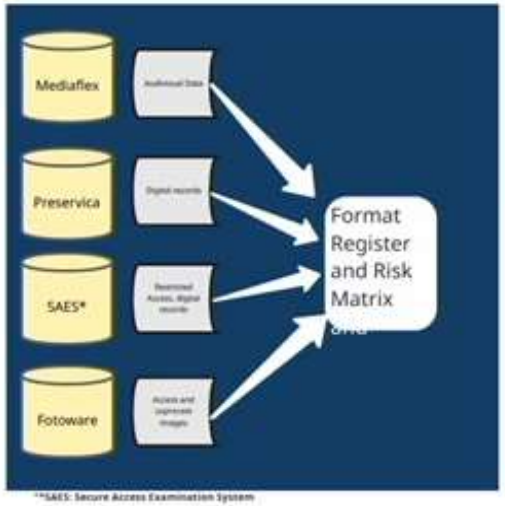
Risk Framework

The National Archives Risk framework initially included a two tiered approach to managing format risk. For all formats in the collection, the format register includes a quick internal triage risk assessment using a series of questions focusing on internal requirements as well as space to include external risk assessments for formats by peer organisations. The second tier was a more complex for detailed technical assessments to support preservation requirements that may include preferred formats for migration for preservation.

Achievements so far

- The file format risk triage tool in the register has been streamlined and the new layout tested for more consistent results
- The pathway to more detailed technical research and preservation planning has been clarified and is being incorporated into ongoing digital archiving workflow uplift
- The data management and reporting aspects of the format register have been separated for improvement in collaboration with ICT systems teams.
- The team has started to link this work with preservation planning for specific electronic record types such as datasets or emails.

Authors:
Elizabeth Long National Archives of Australia
Australia
Elizabeth.long@naa.gov.au
ORCID: 0009-0002-7114-5082
Janet Ma
National Archives of Australia
Australia
Janet.Ma@naa.gov.au
ORCID: 0009-0001-9701-3902
Jonathan Dennis
National Archives of Australia
Australia
Jonathan.dennis@naa.gov.au
ORCID: 0009-0005-3503-6067



**SAES: Secure Access Examination System

File Format Matrix: The Format Matrix was designed for more in-depth evaluation of formats that will support digital archiving decision making. This may include, advice to government agencies, digitisation and preferred formats as well as preservation planning or potential migration for preservation. The matrix currently includes four different sections for assessment: Technical, Documentation, Render Pathway and Essential Characteristics

Key for the digital format risk assessment matrix

File Format Register: this excel document collates data from multiple collection across the archives to assist preservation staff to monitor format prevalence and type as well as assess for areas of high risk that may require action. It contains data about formats including PRONOM FUID and submissions to PRONOM, prevalance in different systems and digital record type to assist preservation planning. Some of this functionality is designed to make up for current system gaps that are being iteratively improved.

Sections of the file format register including category impact and overall risk rating

Future Plans

The review of the current format register has highlighted the importance of system and data connectivity to informed management

- Develop data plan to automate and push format reporting and visibility into archive systems and out of excel
- Build workflows to link and adapt format evaluations to different archiving contexts
- Build internal consistency with file format decision making
- Implement the more detailed risk assessment matrix for identified formats to support preseration planning

This work has highlighted the need to build on the functional model with more detail needed on essential characteristics for common record types
The team want to expand on the use of the matrix to include evaluating potential new formats for migration, digitisation or tranfers advice

Contact: Elizabeth Long Elizabeth.long@naa.gov.au

BnF

iPRES 2025 Posters

iPRES 2025 WELLINGTON AOTEAROA NEW ZEALAND

Collaborative Digital Preservation Policy Revision at Archives Library New Zealand

iPRES 2025 WELLINGTON AOTEAROA NEW ZEALAND

Background

National Library of New Zealand Te Puna Mātauranga o Aotearoa and Archives New Zealand Te Rua Mahara o te Kāwanatanga have a long history of collaboration in digital preservation. This includes shared digital preservation infrastructure, and a joint Digital Preservation Strategy published in 2011.

In 2024, members of both organisations were invited to convene a working group whose goal was to develop a new policy to replace the outdated strategy document, and serve as a foundation for ongoing digital preservation work.

Policy Development Process

Policy Principles

We protect and preserve the digital information in our care.

We have a lifecycle approach to digital preservation.

We champion digital preservation in Aotearoa New Zealand, and support champions of digital preservation across Australasia and the Pacific.

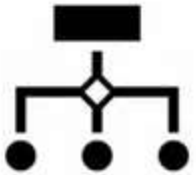
We operate a sustainable digital preservation programme through careful financial stewardship, organisational planning, and close consideration of the environmental impact of our work.

Our digital preservation work supports and is informed by our responsibilities under Te Tiriti o Waitangi.

We recognise that preserving digital information is intrinsically linked with enabling appropriate access, in alignment with current legislation, access restrictions, and treaty obligations.

Next Steps

The National Library and Archives NZ now work under a shared organisational structure with joint leadership. We continue to share and iterate upon the policy based on the feedback we receive from new and existing stakeholders.



The Digital Preservation policy will provide timely guidance to the Digital Preservation Team, who will own policy revisions and any operational impacts.

Conclusions

To remain useful and relevant, policies must be aligned with operational processes and guide future activities. The example of the collaborative policy revision approach taken by the National Library and Archives New Zealand provides a case study for organisations wrestling with their own digital preservation policies.

Te Puna Mātauranga o Aotearoa NATIONAL LIBRARY OF NEW ZEALAND

ARCHIVES NEW ZEALAND

Martin Gengenbach martin.Gengenbach@dia.govt.nz Te Puna Mātauranga o Aotearoa National Library of New Zealand | Digital Preservation Policy Revision Working Group: Nicola Caldwell, Rhonda Grantham, Jan Hutar, Blanche Joselin, Carly Lenz, Val Love, Charlotte McGillin, Joshua Ng

{BnF

iPRES 2025 Posters

Standard Transfer Metadata for Born Digital Transfers

Jan Hutař, Archives New Zealand, jan.hutar@dia.govt.nz

Born digital transfers

- Public Records Act 2005 – government agencies to transfer records in ANY form
- Urgent digital transfers accepted from 2006 - no means of processing or actively preserving them, just secure storage
- 2010-2013 Government Digital Archive Programme (GDAP), aimed to deliver:
  - new digital preservation system ☑
  - ability to process digital transfers ☒

Transfer requirements

- GDAP closed early but the need for transfer process ability did not disappear
- Built minimal requirement workflow and process (python scripts, guidelines, etc.) > first "official" transfer in 2015 and all subsequent transfers
- Requirements for agencies were simple (we thought), had to change to bare minimum over time
  - transfer files "as they are" in any format; do not migrate formats before transfer
  - metadata in machine-readable form (ideally CSV), all the metadata fields

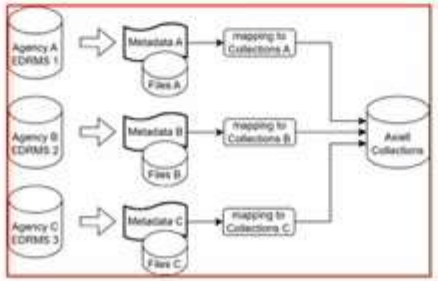
Every transfer is a project

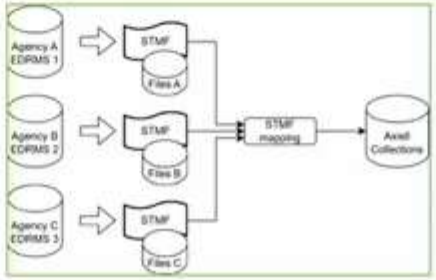
- Lack of metadata standard is an issue
  - field naming inconsistencies
  - content issues
  - format and encoding quirks
- For each transfer we need to:
  - analyse the metadata fields
  - map them to fields used in our archival description system

STMF

Standard Transfer Metadata File

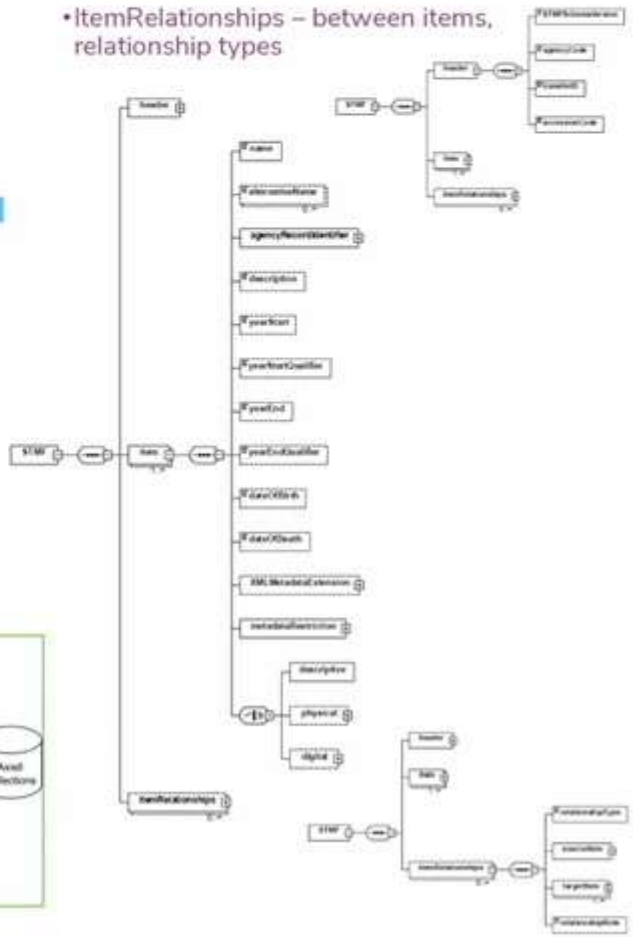
- Prescribing how the metadata for born-digital transfers must look
- XSD and spreadsheet template + detailed guidelines
- Agencies will use those to export metadata from their systems (vendor cooperation will be needed)
- Export templates already exist in some EDRMS, e.g. NAA (National Archives of Australia), VEO (Victoria, Australia), etc.
- We plan to provide supplementary tools like an STMF validator





STMF structure

- Header – details about the transfer
- Item – item/records descriptive and technical metadata
- ItemRelationships – between items, relationship types



iPRES 2025 Posters

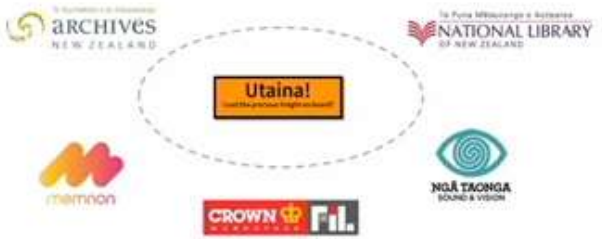
iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

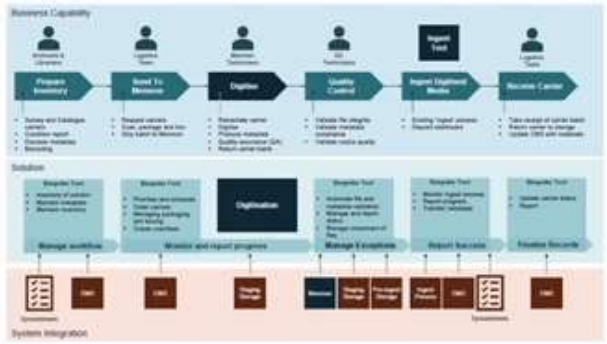
Utaina

Connecting Aotearoa Through Digitised Sounds and Stories

Utaina

A collaborative project of Archives New Zealand (ArchivesNZ), the National Library of New Zealand (NLNZ) and Ngā Taonga Sound & Vision (NTSV) to digitally preserve more than 400,000 items of Aotearoa New Zealand's Crown-owned at-risk audiovisual magnetic media.





Balancing Complexities and Scale

Scale

ArchivesNZ and NLNZ - Digitised over **70,000** items (73,000 playback hours) across 13 formats within 5 years.

Preservation Format & Digital Storage

FFV1/MKV lossless compression as preservation video format to counter digital storage cost pressure. Storage size savings up to **50%**.

Collection Discovery & Description

Fresh collections & holdings rediscovery uncovered over **20,000** previously undescribed and inaccessible items.

Logistics

Movement and tracking of over **1,000** items weekly at every stage from packing to digitisation and back on shelves.

Quality Control

Strategic and scalable QC—automated validation and statistical sampling—enabled us to meet preservation standards within strict workflow deadlines.

Impact

Digitised audiovisual items now accessible virtually and internationally.

Some items previously found with mould now cleaned and digitised.

Culturally significant audiovisual items are rediscovered and made accessible.

Reconnecting researchers to collections made by their family, or on labels that are defunct and previously not available anywhere.

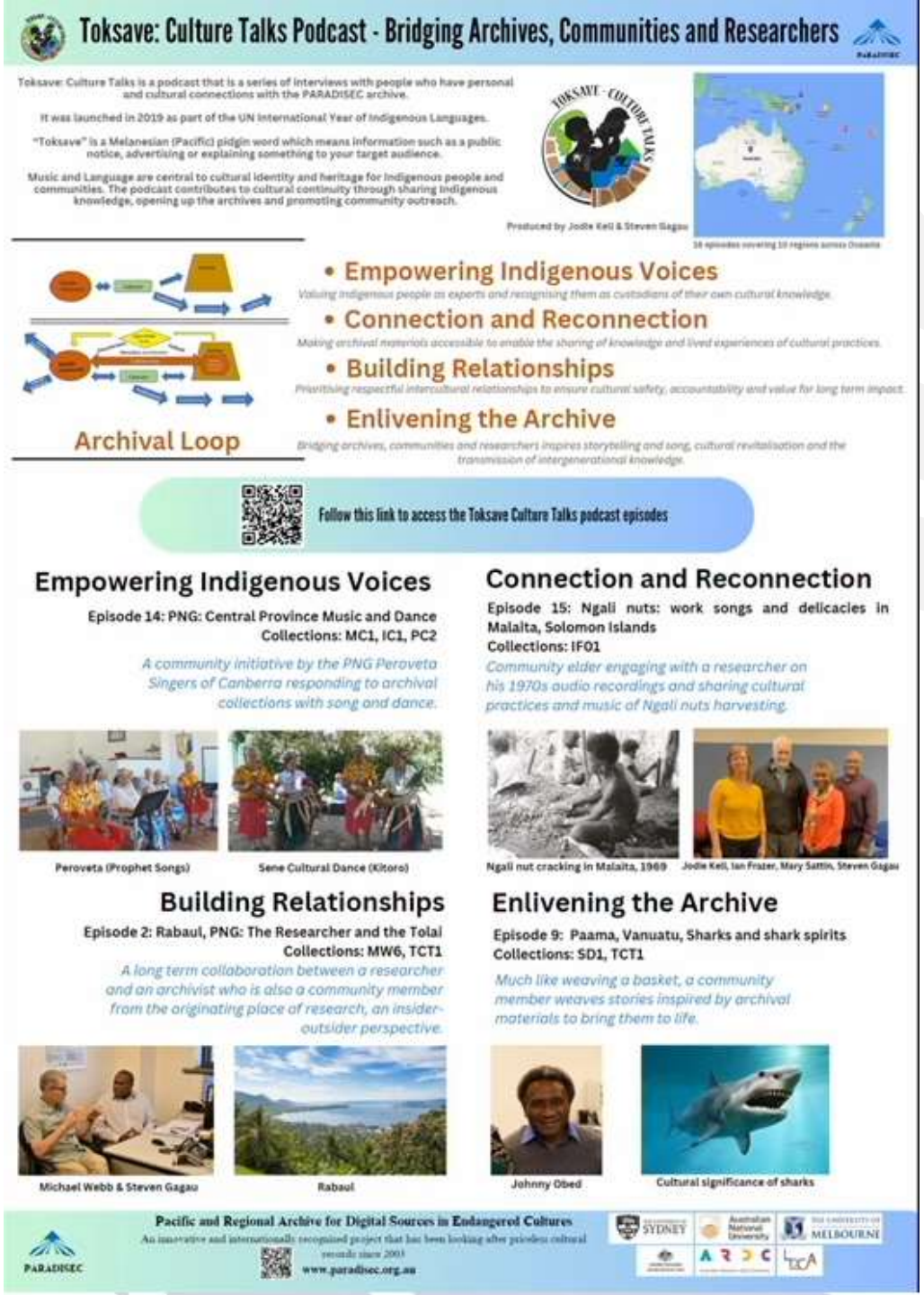
Preserving the nation's memory and record of our Government - snapshots of how the state once imagined its relationship with people and the land.

Poster created for iPRES 2025 by Joshua Ng, Cynthia Wu and Zak Agrabrite
utainaproject@dia.govt.nz

BnF

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

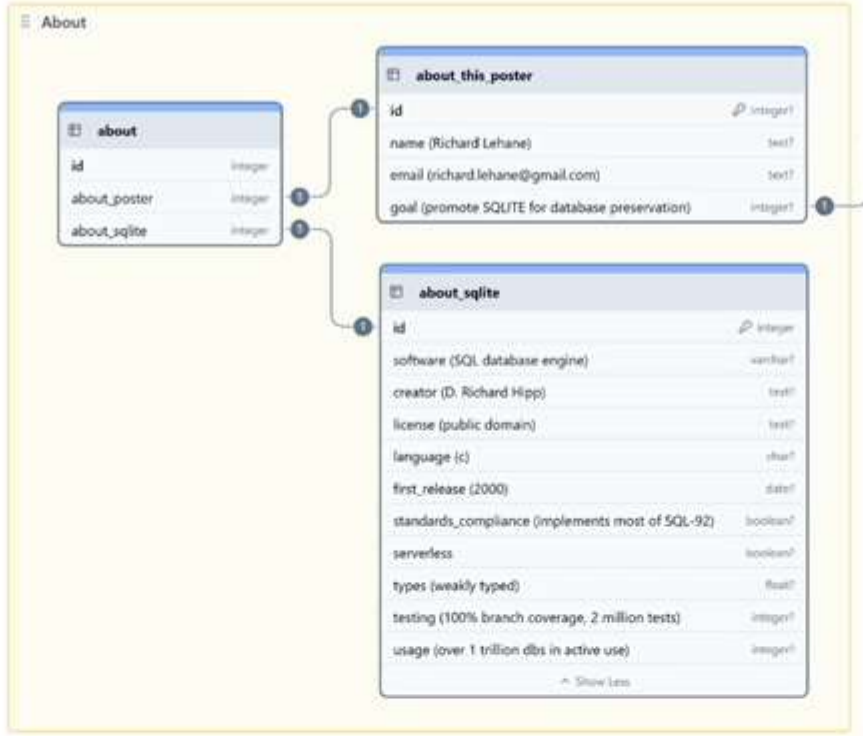


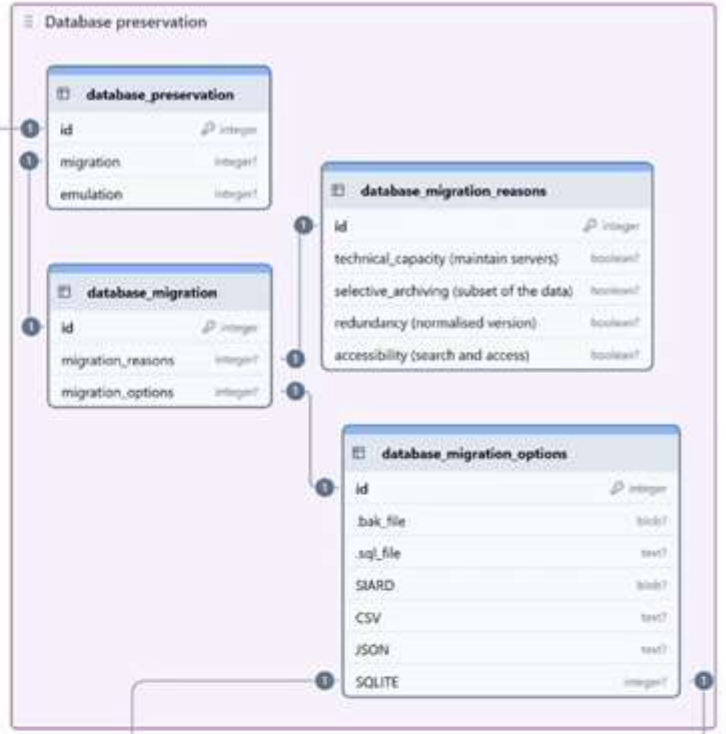
iPRES 2025 Posters

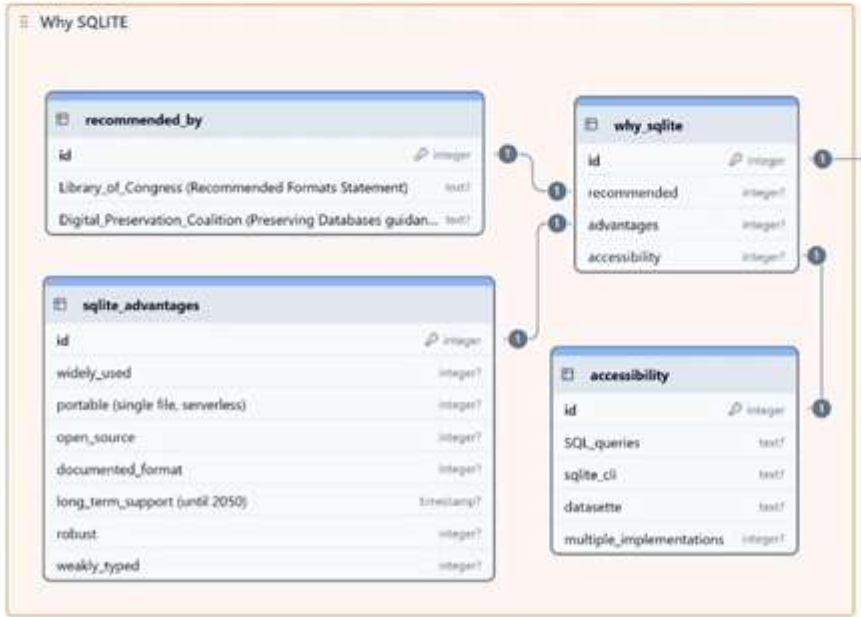
SQLite for database preservation

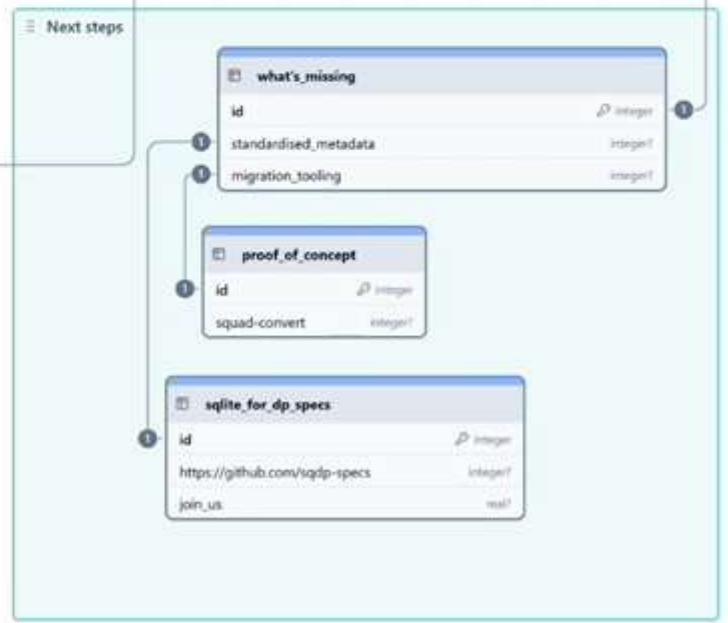
Richard Lehane
richard.lehane@gmail.com
https://github.com/sqdp-specs

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND









{BnF

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

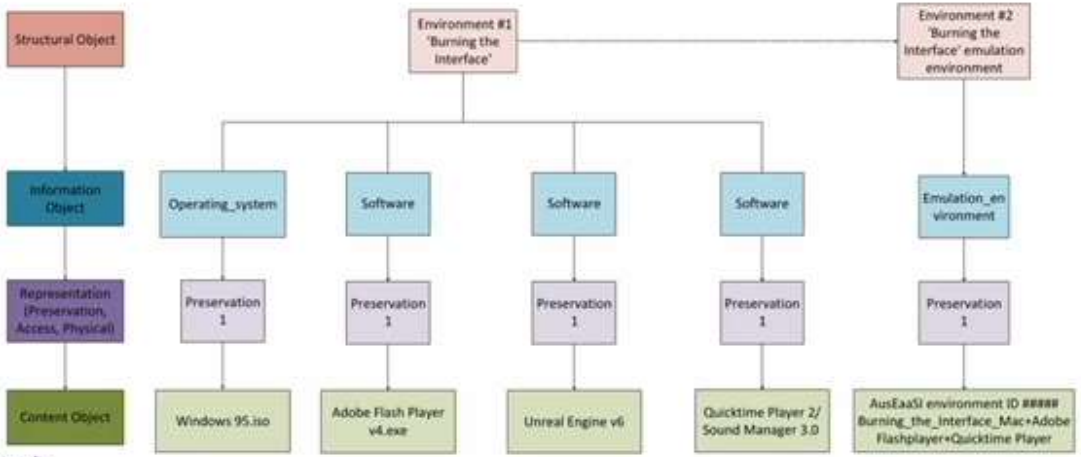
Art Gallery NSW

How the Art Gallery of New South Wales Represents software and hardware environments within its digital preservation system

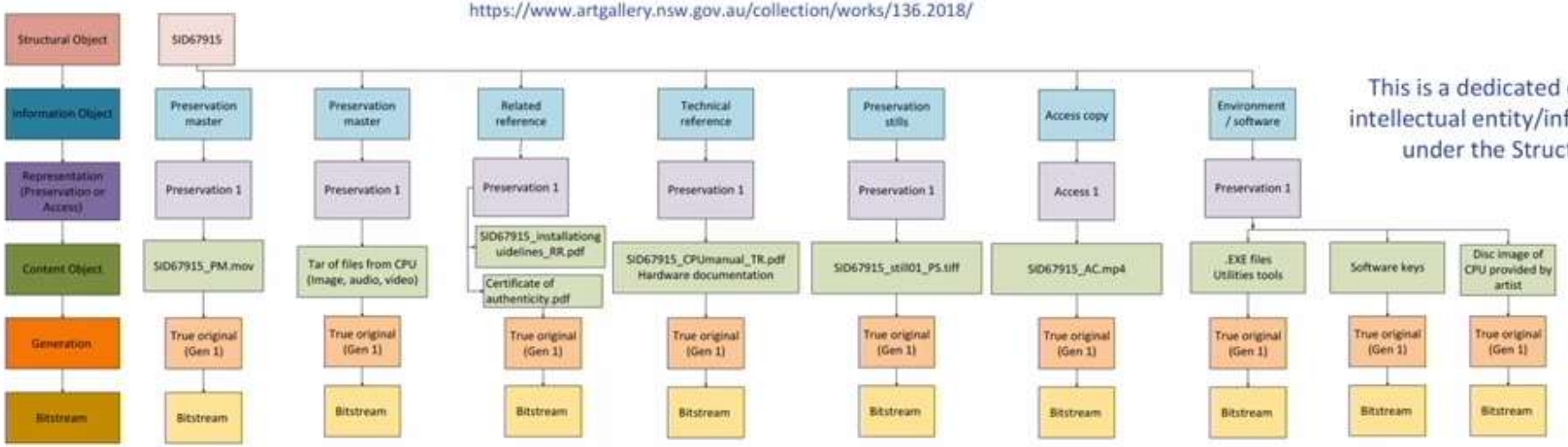
Mapping PREMIS environmental entities to the AGNSW and Preservica data models

This IE is linked to artworks requiring this environment for access.

Emulation environments can also be added as a representation to the artwork



More complex example
- Algorythmic video work
https://www.artgallery.nsw.gov.au/collection/works/136.2018/



This is a dedicated environmental intellectual entity/information object under the Structural object

Contributors:
Joanna Fleming: Joanna.Fleming@ag.nsw.gov.au, Ziggy Potts: Ziggy.Potts@ag.nsw.gov.au & Bec Barnott-Clement: Rebecca.Barnott-Clement@ag.nsw.gov.au

BnF

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Bridging Communities and Archives

Preserving digital Documents with the Citizen Archive Platform

The **Citizen Archive Platform** is a web-based tool that facilitates the structured transfer of born-digital data from individuals, communities and cultural initiatives to archives for long-term storage. The goal is to simplify and standardise the submission process.

Technical details

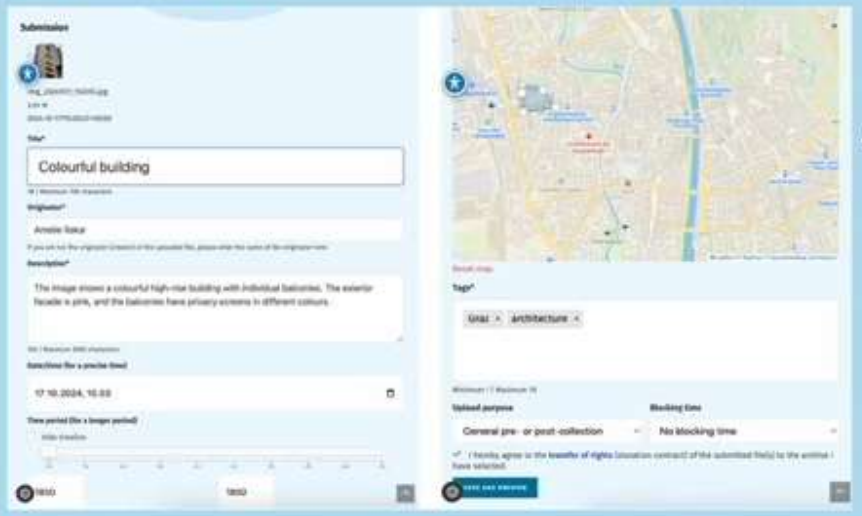
- CMS Word Press
- SQL Database
- Specialised Plugin for data processing

Data processing

- Upload-interface
- Automatic format validation and virus scan
- Technical metadate extraction

Submission Information Package
with files + all technical and content metadata

Metadata form



User: Registration → Data upload → Metadata entry → Submission

Archivist: Appraisal → Metadata entry → Download SIP → Ingest: Digital Archive

Good to know

- Standards-compliant (OAIS, METS, EAD)
- Technical and content metadata are combined in one XML
- Legal issues (copyright, blocking time and contract) are clarified during the submission process
- Open source: Other archives can use the platform; Publication via GitHub at the end of 2025

Amelie Rakar
amelie.rakar@stadt.graz.at

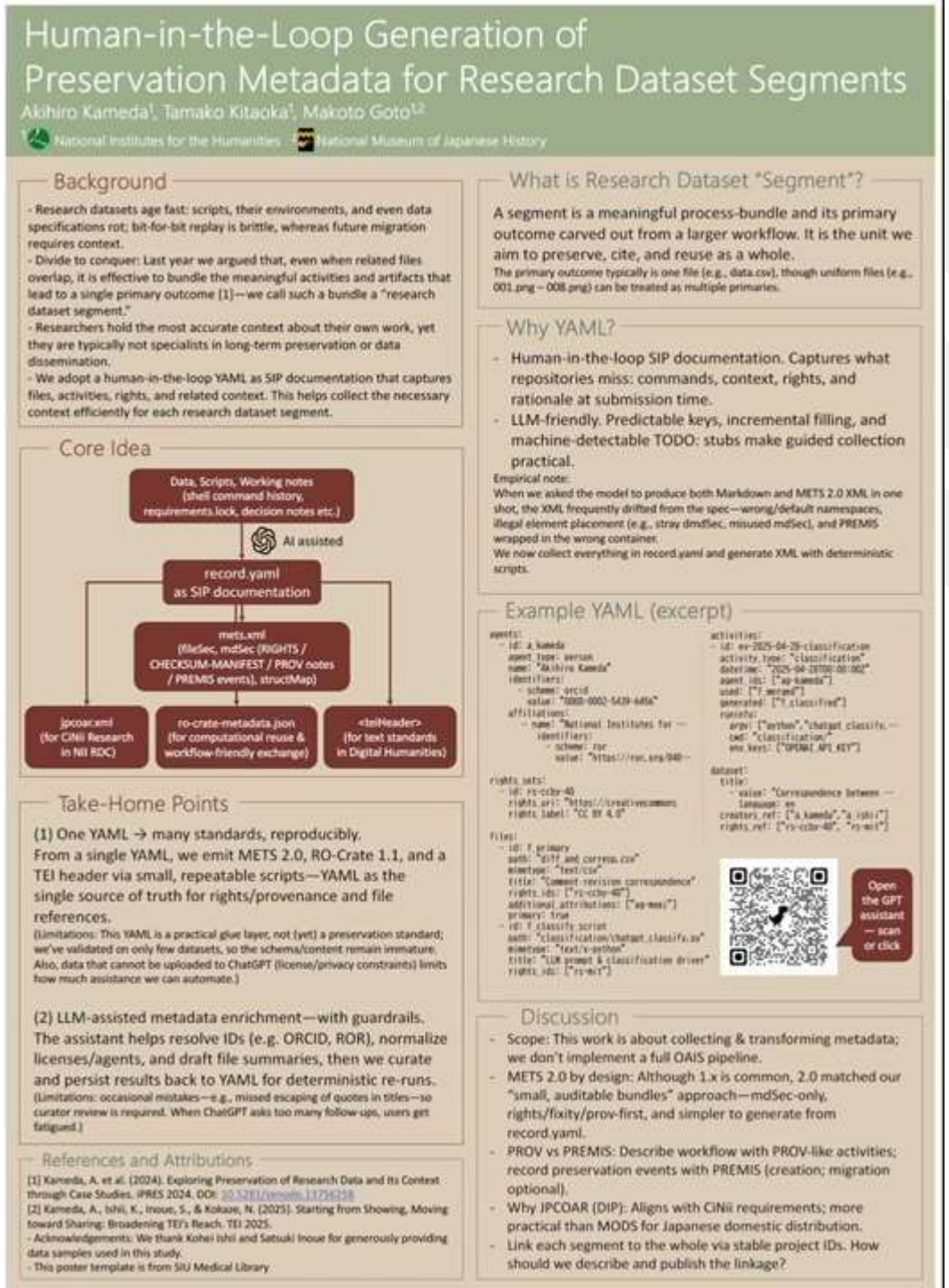
citizen archive platform

Co-funded by the European Union

BnF

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND



Human-in-the-Loop Generation of Preservation Metadata for Research Dataset Segments

Akihiro Kameda[1], Tamako Kitaoka[1], Makoto Goto[1,2]
[1] National Institutes for the Humanities [2] National Museum of Japanese History

Background

- Research datasets age fast: scripts, their environments, and even data specifications rot; bit-for-bit replay is brittle, whereas future migration requires context.
- Divide to conquer: Last year we argued that, even when related files overlap, it is effective to bundle the meaningful activities and artifacts that lead to a single primary outcome [1]—we call such a bundle a "research dataset segment."
- Researchers hold the most accurate context about their own work, yet they are typically not specialists in long-term preservation or data dissemination.
- We adopt a human-in-the-loop YAML as SIP documentation that captures files, activities, rights, and related context. This helps collect the necessary context efficiently for each research dataset segment.

Core Idea

Take-Home Points

(1) One YAML → many standards, reproducibly.
From a single YAML, we emit METS 2.0, RO-Crate 1.1, and a TEI header via small, repeatable scripts—YAML as the single source of truth for rights/provenance and file references.
(Limitations: This YAML is a practical glue layer, not (yet) a preservation standard; we've validated on only few datasets, so the schema/content remain immature. Also, data that cannot be uploaded to ChatGPT (license/privacy constraints) limits how much assistance we can automate.)

(2) LLM-assisted metadata enrichment—with guardrails.
The assistant helps resolve IDs (e.g. ORCID, ROR), normalize licenses/agents, and draft file summaries, then we curate and persist results back to YAML for deterministic re-runs.
(Limitations: occasional mistakes—e.g., missed escaping of quotes in titles—so curator review is required. When ChatGPT asks too many follow-ups, users get fatigued.)

References and Attributions

[1] Kameda, A. et al. (2024). Exploring Preservation of Research Data and Its Context through Case Studies. iPRES 2024. DOI: 10.5281/zenodo.13754258
[2] Kameda, A., Ishii, K., Inoue, S., & Kokaze, N. (2025). Starting from Showing, Moving toward Sharing: Broadening TEI's Reach. TEI 2025.
- Acknowledgements: We thank Kohei Ishii and Satsuki Inoue for generously providing data samples used in this study.
- This poster template is from SIU Medical Library

What is Research Dataset "Segment"?

A segment is a meaningful process-bundle and its primary outcome carved out from a larger workflow. It is the unit we aim to preserve, cite, and reuse as a whole.
The primary outcome typically is one file (e.g., data.csv), though uniform files (e.g., 001.png – 008.png) can be treated as multiple primaries.

Why YAML?

- Human-in-the-loop SIP documentation. Captures what repositories miss: commands, context, rights, and rationale at submission time.
- LLM-friendly. Predictable keys, incremental filling, and machine-detectable TODO: stubs make guided collection practical.

Empirical note:
When we asked the model to produce both Markdown and METS 2.0 XML in one shot, the XML frequently drifted from the spec—wrong/default namespaces, illegal element placement (e.g., stray dmdSec, misused mdSec), and PREMIS wrapped in the wrong container.
We now collect everything in record.yaml and generate XML with deterministic scripts.

Example YAML (excerpt)

```
agents:
  - id: a_kameda
    agent_type: person
    name: "Akihiro Kameda"
    identifiers:
      - scheme: orcid
        value: "0000-0002-5439-645a"
    affiliations:
      - name: "National Institutes for ..."
        identifiers:
          - scheme: ror
            value: "https://ror.org/048..."

rights_sets:
  - id: rs-ccby-4.0
    rights_uri: "https://creativecommons
    rights_label: "CC BY 4.0"

files:
  - id: f_primary
    path: "diff_and_correspo.csv"
    mimetype: "text/csv"
    title: "Comment-revision correspondence"
    rights_ids: ["rs-ccby-4.0"]
    additional_attributes: ["ap-kmd"]
    primary: true
  - id: f_classify_script
    path: "classification/chatgpt_classify.py"
    mimetype: "text/x-python"
    title: "LLM prompt & classification driver"
    rights_ids: ["rs-mit"]

activities:
  - id: ev-2025-04-29-classification
    activity_type: "classification"
    datetime: "2025-04-28T00:00:00Z"
    agent_ids: ["ap-kameda"]
    used: ["f_merged"]
    generated: ["f_classified"]
    runinfo:
      argv: ["python","chatgpt_classify.py"]
      cwd: "classification/"
      env_keys: ["OPENAI_API_KEY"]

dataset:
  title:
    - value: "Correspondence between ..."
      language: en
  creators_ref: ["a_kameda","a_ishii"]
  rights_ref: ["rs-ccby-4.0", "rs-mit"]
```

Open the GPT assistant — scan or click

Discussion

- Scope: This work is about collecting & transforming metadata; we don't implement a full OAIS pipeline.
- METS 2.0 by design: Although 1.x is common, 2.0 matched our "small, auditable bundles" approach—mdSec-only, rights/fixity/prov-first, and simpler to generate from record.yaml.
- PROV vs PREMIS: Describe workflow with PROV-like activities; record preservation events with PREMIS (creation; migration optional).
- Why JPCOAR (DIP): Aligns with CiNii requirements; more practical than MODS for Japanese domestic distribution.
- Link each segment to the whole via stable project IDs. How should we describe and publish the linkage?

BnF

iPRES 2025 Posters

HISTORICAL SOCIETIES AND DIGITAL PRESERVATION CHALLENGES

BACKGROUND

Grid Heritage collects historic equipment, film footage, photographs and documents relating to electricity transmission in New Zealand.

THE CHALLENGE

How will historical societies like Grid Heritage pivot from physical to digital preservation?

CONCLUSION

By proactively preparing to receive digital records, Grid Heritage can safeguard the rich history of New Zealand's electricity transmission for future generations.

KEY QUESTIONS

How can small historical societies prepare themselves for receiving digital content and preserving digital formats?

What resources are available for them?

How does an amateur historical society obtain the necessary expertise that is required to properly preserve its historical records?

TRANSPOWER

Gunilla Jensen, Transpower New Zealand: Gunilla.Jensen@transpower.co.nz
Michael Williams, Transpower New Zealand: Michael.Williams@transpower.co.nz

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

LESSONS LEARNED

AGNSW now produce detailed signal flow diagrams—using web-based diagramming tool draw.io—as part of its suite of Conservation documentation for complex works.

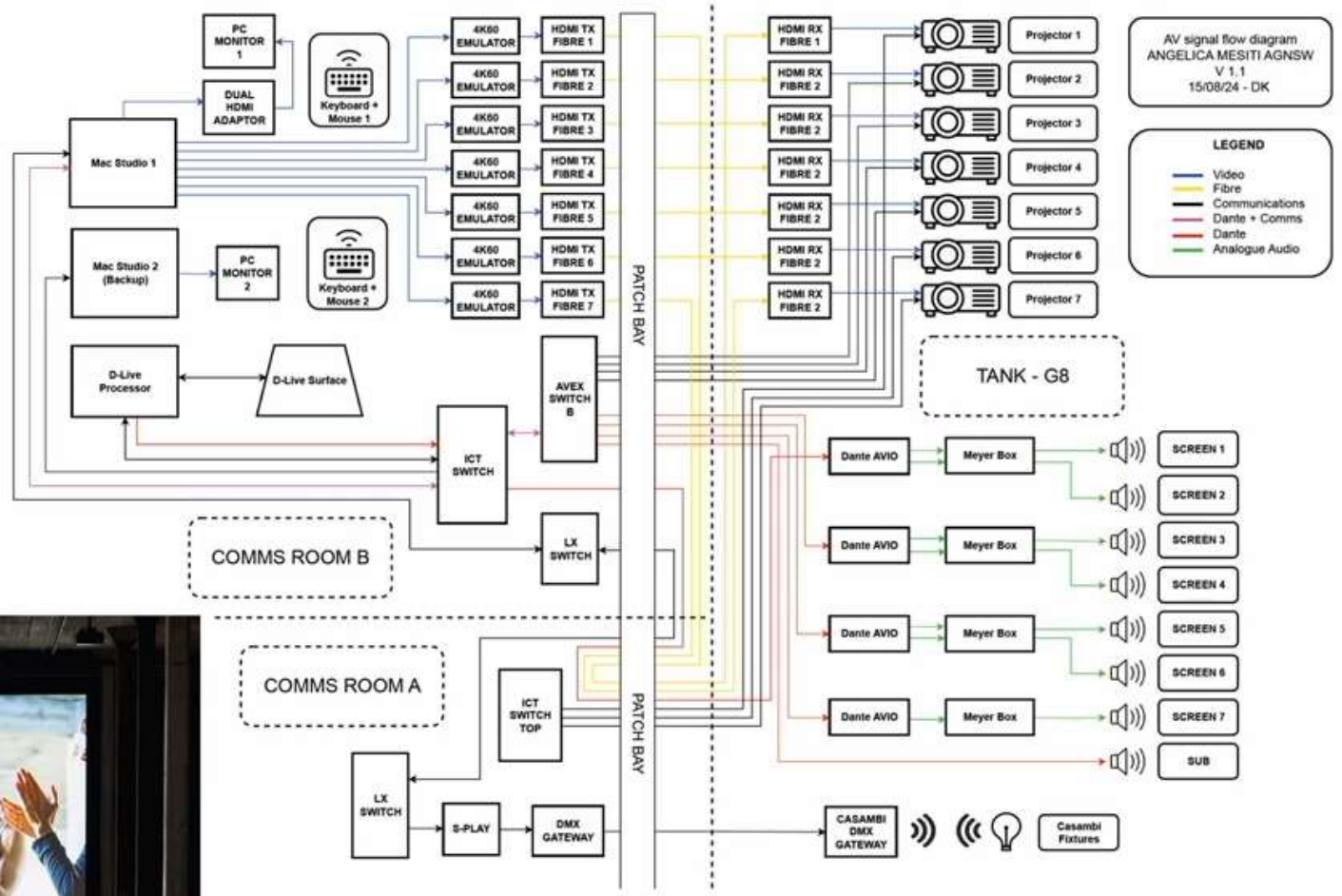
Pictured here is The Rites of When (2024) by artist Angelica Mesiti. A rich visual and sonic installation commissioned for the Nelson Packer Tank gallery, a former wartime oil bunker beneath the Naala Badu building. The work features seven monolithic vertical screens arranged in a circle within the 2,200-square-metre space.

The correspondig diagram illustrates how the artwork is streamed from two Apple Mac Studios networked together and running QLab. Synchronization is achieved via audio timecode, with Mac Studio 1 acting as the master timecode generator. The timecode signal is sent to the dLive mixer and then returned to both Mac Studios, ensuring precise, synchronized playback across both systems.

The dLive mixer also receives the audio from Mac Studio 1, routing Dante audio via the switch to Meyer speakers and subwoofer. Video outputs pass through 4K60 emulators to HDMI transmitters and receivers, driving seven projectors displaying visuals on microperforated screens. Audio travels over Dante to AVIO interfaces connected to Meyer box speakers embedded in the screens.

The intricate systems behind GROUNDLOOP and The Rites of When demonstrate that preserving media artworks is as much about understanding infrastructure as it is about safeguarding the artist's Master files.

Customised documentation—like signal flow diagrams, network maps, and codec specifications—is vital for installation, troubleshooting, and upholding the artist's vision. By capturing these details with precision, we ensure that time-based artworks can be authentically re-experienced far into the future.



Above: 'The Rites of When' signal flow diagram by Daniel Kilp.

Left: Installation view of Angelica Mesiti 'The Rites of When', 2024, 7-channel digital video installation, colour, sound, approx 20 min, collection of the artist, commissioned by the Art Gallery of New South Wales for the Nelson Packer Tank, 2024 © Angelica Mesiti, photo © Art Gallery of New South Wales, Jenni Carter

Art Gallery NSW

{BnF

iPRES 2025 Posters

iPRES 2025 WELLINGTON
AOTEAROA NEW ZEALAND

Preservica

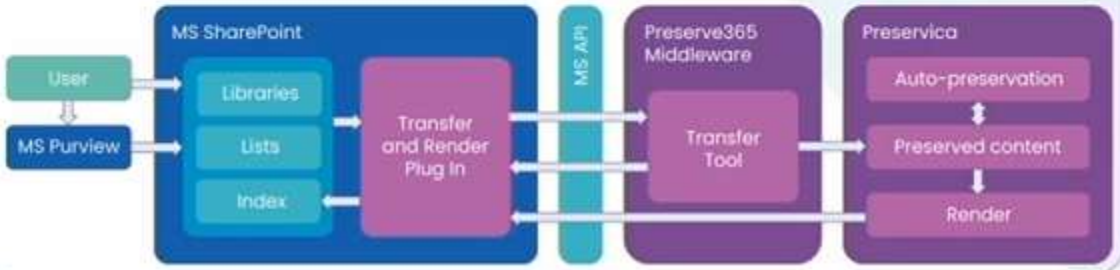
The integration of Digital Preservation into SharePoint Online

SharePoint information

- ✓ Library: File hierarchy, one file per record, user and system metadata
- ✓ Lists: Emphasis on tabular metadata, zero to many attachments, folder hierarchy
- ✗ Sites: Website based data presentation

Alternate Integration Models

- ✗ Web archiving: Easy to do for Sites but lossy for Libraries and Lists
- ✗ Extract, Transform and Load: End of life transfer using export tools but no SharePoint user control
- ✗ Browse and extract with API & external application: Can't be done inside SharePoint
- ✓ Tagged transfer using API: Use user metadata or retention or sensitivity labels to extract and transfer data
- ✓ Transfer plug-in: Users can configure and transfer data inside SharePoint
- ✓ Search and Render plug-in: User can search and render preserved content in SharePoint



Preserve365® Key Features

- Middleware application presents a GUI within Microsoft and transfers data to and from SharePoint and Preservica
- Transfer plug-in allows users to select data from Libraries or Lists and move or copy

- Background application searches for items tagged with specific retention labels, sensitivity labels or metadata fields which are moved or copied for Preservation
- Compatible with Copilot, Purview, PowerShell
- Preserved content is indexed in SharePoint to allow search from within the Microsoft world
- A render plug-in allows users to render Preserved content within SharePoint

Challenges

- Significant software development costs
- Microsoft API extensions and updates, but Microsoft were a great partner
- Large data volumes
- Applications need elevated permissions
- Some MS features require expensive licenses
- Transfer validation limited by changes to content on exit from SharePoint

Preserving SharePoint Records in Preservica

- Multi-part asset comprising the metadata in JSON, content files, complex metadata files (images, rich text)
- JSON contains all SharePoint user and system metadata for the record and is fully preserved
- SharePoint metadata also in XML for search
- Record hierarchy is maintained
- Middleware transfers complete packages which get the full DP ingest workflow
- Formats migrated according to Preservation Plan which can be changed post-ingest
- Preserved items are indexed within SharePoint index with original access control
- Specialist render of complete Library or List item within Preservica and SharePoint with all metadata (user defined and system)

What Next (all work in progress)

- Further integration with Copilot to drive transfer and enable content use
- Preserving multiple versions of records
- Preserve emails from Outlook
- Preserve chat, meetings & channels in Teams
- Preserve files from OneDrive

BnF