

From legacy to next generation: a story of collaboration to push the boundaries of the open source Haplo repository from Cayuse

This article describes the development of the Haplo standards-based, open source repository software, from Cayuse, that meets the findable, accessible, interoperable and reusable (FAIR) principles, and which captures all research, regardless of what it is, with a focus on prioritizing the capture of 'practice research' – '... an umbrella term that describes all manners of research where practice is the significant method of research conveyed in a research output.' in the arts and architecture. This research has been neglected by the repository landscape and surrounding discoverability infrastructure, which has traditionally focused on text-based research publications in the STEM disciplines where there is a policy mandate (and funding) for open access. As practice research has not been captured effectively in repositories, it has not been possible for it to be preserved for long-term access via appropriate integrated digital preservation solutions. This story of collaboration between the University of Westminster and Haplo (now Cayuse), puts researchers at the centre of development, using a co-design approach, while ensuring the Research and Scholarly Communications team (then based within the Library and Archives Service) at the University were driving this work in alignment with sector-wide standards. The role of user engagement, advocacy and inclusive policy development is highlighted and illustrates that this underpins, and is crucial to, successful software development. While the successes are documented and celebrated, the challenges are acknowledged and the lessons learned are shared.

Keywords

Haplo; Cayuse; digital repositories; open source software; practice research; non-text outputs



JENNY EVANS

Research
Environment and
Scholarly
Communications
Lead
University of
Westminster



NINA WATTS

Repository and Open
Access Manager
University of
Westminster



TAYLOR MUDD

Senior Software
Engineer
Cayuse LLC



TOM RENNER

(Previously Haplo
Services Ltd.)

Background

The University of Westminster and Haplo (Note 1) began working together in 2012 on a project to address the research data management requirements of research funders.¹ Led by the Information Services directorate (a converged Library and IT support service),

2 researchers and support staff were embedded in the development process from the beginning. Initial interviews revealed research-related processes, particularly those concerning the PhD student journey and research ethics applications, were the highest priority for the community.² This led to the development of the Haplo research information management system, which now includes pre-award processes (including a costing tool) and post-award processes, repository, PhD progression and ethics processes. This is known collectively within the University as the Virtual Research Environment (VRE).

WestminsterResearch, the University’s institutional repository, pre-dates the development of the Virtual Research Environment, with a mediated deposit service using EPrints,³ launched in 2006–07. 2014 brought the implementation of a hybrid solution which saw a VRE (Haplo)-EPrints integration, with the VRE used as the user interface and a data feed to EPrints for open access (OA) discoverability. This improved user experience allowed a change from mediated to self-deposit (partly attributable to the Research Excellence Framework (REF) 2021 Open Access (OA) policy⁴ launch at the same time – REF is the UK’s periodic national assessment of the quality of research across disciplines). There was a subsequent increase (illustrated below) of the total number, and the number of outputs including full text in WestminsterResearch (Figure 1) and the cumulative percentage of full text over time (Figure 2). However, the Haplo-EPrints integration was unstable, with any update to EPrints breaking the integration. This approach also prevented us from benefiting from the flexibility of the Haplo repository architecture, as metadata entered into the Haplo user interface had to map to EPrints.

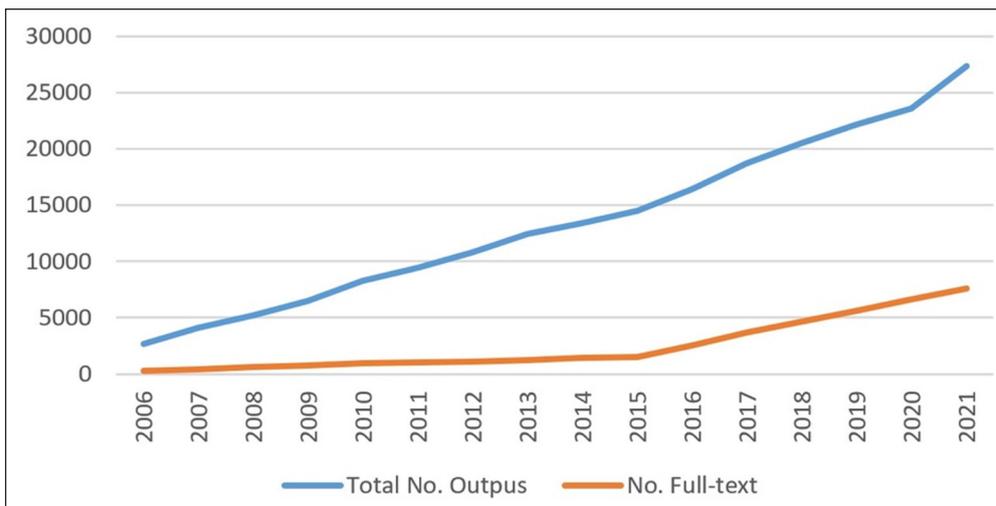


Figure 1. Number of outputs included in WestminsterResearch (1 Jan 2006 – 31 Dec 2021)

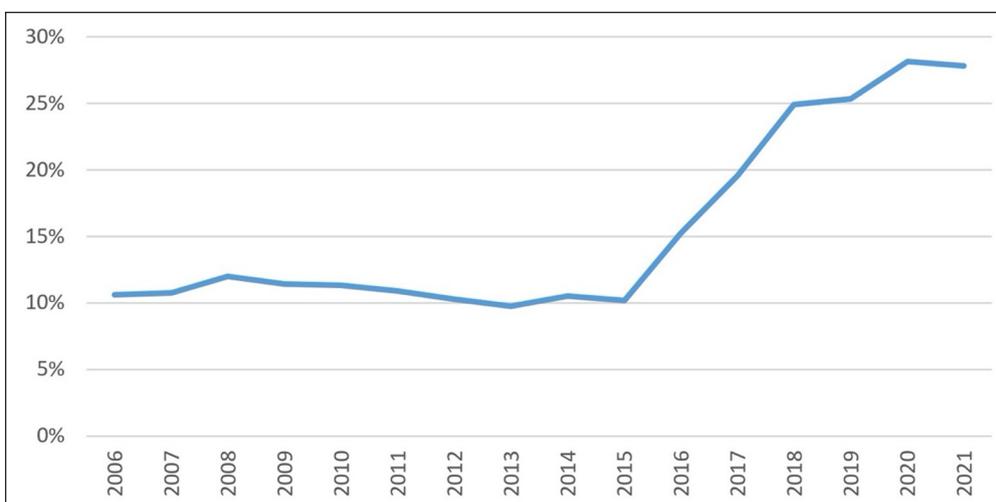


Figure 2. Cumulative % Full-Text (1 Jan 2006 – 31 Dec 2021, Source: WestminsterResearch)

3 Making the case for an all-Haplo repository: strategic investment

EPrints was trusted, it had underpinned the submission of research outputs to REF2014, had an active user community and a proven record of discoverability by search engines. A benefits, risks and mitigations exercise carried out in 2017 focused on requirements, standards and interoperability. The benefits were clear: all research outputs are in one repository including datasets and practice research in the arts and architecture and the cost savings achieved by having just one repository subscription. The identified risks of staying with the hybrid solution (not meeting funder research data management requirements, not capturing the entirety of the University of Westminster research outputs, not having a user community) were higher than the perceived risks of being a development partner and moving (putting 'all our eggs in one basket'). These risks could be mitigated by having an exit strategy in place.

'The Haplo repository is built on open source repository technology'

Haplo repository technology

The Haplo repository is built on open source repository technology and has several layers (Figure 3). The platform layer manages security, permissions and version control.

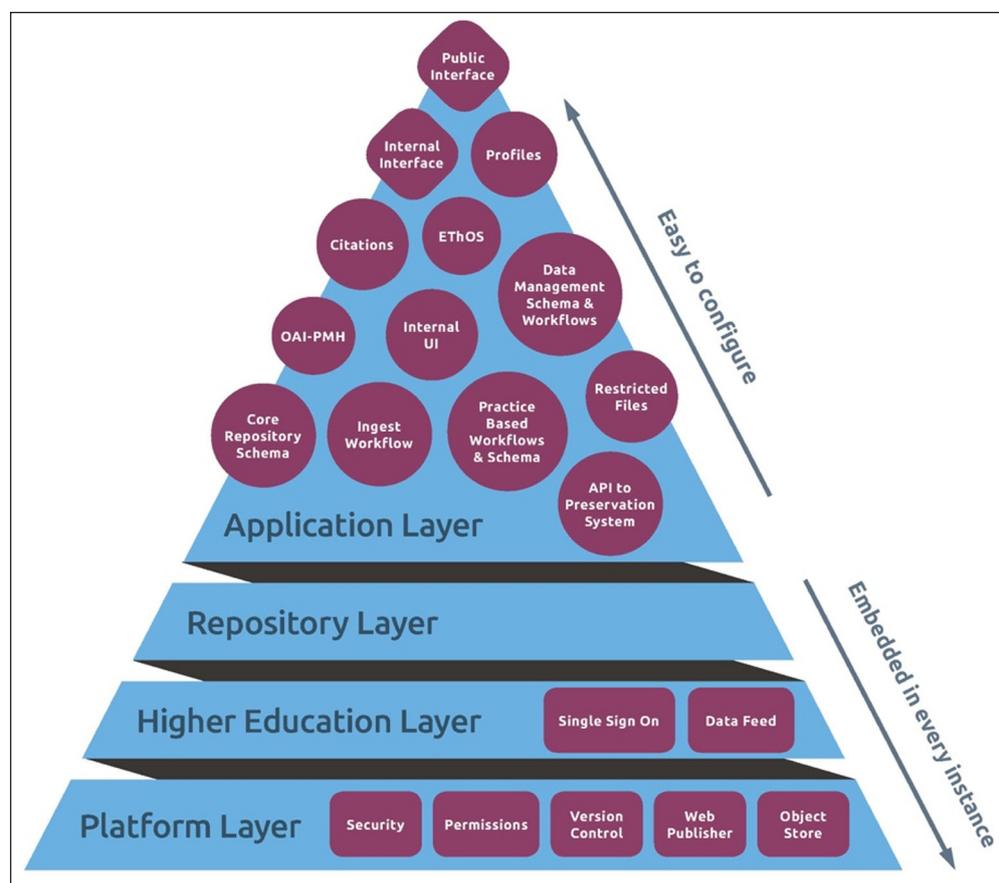


Figure 3. Haplo repository architecture

Records are stored in a linked object datastore, which can hold records of any type, and includes unique records for journals, publishers, funders and projects. In addition, each record is assigned an identifier, typically a five-character alphanumeric sequence that is unique within a given system. These unique identifiers facilitate the primary data type within the Haplo system, a link, which also helps the system meet linked data principles.⁵ The power of these links is leveraged within the repository as the ability to link records together within the system, and draw information from each, allows for complex records such as repository outputs to be simply and accurately represented with little data re-entry or duplication. For example, the record for an output will include a link to the unique record for the journal – which may provide access to the relevant ISSN and publisher.

- 4 By navigating this linked datastore, complex graph queries can provide sophisticated insights into the data. Entering data as a link to an existing record preferentially throughout the application minimizes the amount of duplicate or misspelled data. A key benefit is to uniquely identify people within the system, ensuring correct attribution for their role on a research output.

In line with linked data principles and search engine optimization best practice (meaning better discoverability), the repository pages are written with clear semantic mark-up, embedded metatags, canonical links for machine-readable data and meet responsive design principles. It meets the findable, accessible, interoperable and reusable (FAIR)⁶ principles with outputs assigned DOIs or handles, implements the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)⁷ to enable discovery by standard industry discovery tools, and an authentication process is used to manage access to datasets. Metadata and persistent identifier standards, including Dublin Core, OpenAIRE, Crossref and DataCite, are applied to enable interoperability, each output is described appropriately and a reuse licence applied.

'A key benefit is to uniquely identify people within the system, ensuring correct attribution'

The repository software enables a public interface that is easily changed, offering unlimited control over the visual design by utilizing customizable layouts, templating of pages and allowing institutional-specific style sheets to be used. This flexibility affords many benefits: the ability to build a bespoke public interface for an institution to provide a consistent look and feel across institutional sites, increased discoverability on the web to further the reach of outputs within the repository and meeting accessibility best practice.

'The repository software enables a public interface that is easily changed'

The development process, hosting and software is ISO/IEC 27001:2013⁸ certified, making it suitable for handling sensitive records such as datasets. Access to restricted files and records is managed using a fine-grained, role-based access control system. Permissions rules can be applied to individual fields on records, using defined access levels to control access to sensitive files.

Haplo operates a 'software-as-a-service' business model, providing hosting, development and implementation services for their software. The repository software is released as open source, available for anyone to download, install, run and extend. The benefits of this software are therefore shared more widely, actively improved, bugs are fixed and new features are added to keep up with the community's requirements. The resulting distributed development and opportunity for implementation of the solution can extend the solution's reach and assist with bringing in sales. This mutually beneficial arrangement is typical of successful open source software projects, and key to long term viability.

Methodology

User engagement: an agile approach to development

Using a co-design approach, we included academic stakeholders and formed a specialist 'practice research in the arts and architecture' advisory group, with whom we worked throughout the project. Understanding their drivers was a priority and their approach to the submission of portfolios (collections) to REF2021 provided a useful structure. We met regularly with the advisory group using tools including process maps (effective for gaining consensus on how a workflow should look). Template examples helped identify enhancements to better reflect what the research looks like, who contributes to its creation and what would make it more user-friendly. We demonstrated various iterations of test systems and recorded their feedback to inform development. A key part of this engagement was going to the research community – holding meetings at their campuses, joining their away days to obtain feedback – which contributed to the creation of a collaborative environment.

Requirements gathering

The Research and Scholarly Communications team started with a long list of requirements, informed by its experience about what was and was not working for researchers and administrators, and refined it in consultation with colleagues within the University and two other institutions considering a move to Haplo. We discussed with Haplo how the technology could solve these issues by building a standards-based repository that harnessed flexible architecture to capture all research output types and meet funder requirements.

Practice research in arts and architecture

The Centre for Research in Education in Arts and Media (CREAM) at the University is a world-leading centre and pioneer in practice-based, critical, theoretical and historical research in the broad areas of art, creative and interdisciplinary practice. The Making and Practice research group is engaged in creative practices within architecture. Discussions with these researchers enabled us to understand what a practice research output would ideally look like. What we learned was that each individual research output could be a publication, research dataset or a non-text (non-traditional) output. These outputs then needed to be connected together into a collection (portfolio) and the underlying research methods needed to be documented by a narrative. For example, an artefact may be exhibited multiple times, have a catalogue published, then perhaps a journal article, or a report, would be published at a later stage. Vocabulary was important, words such as 'author' and 'abstract' were not seen as relevant, the roles of collaborators significant, for example curator, producer, set designer. One unified repository meant researchers did not have to think about 'outputs' separately to the 'data' being created.

Previous UK-based projects customizing EPrints for arts research⁹ had done much to address the challenges, for example KULTUR (2007–2009)¹⁰ produced an EPrints plug-in for arts-based institutional repositories, which was adopted by a number of institutions and led to further work including Kultivate¹¹ and Defiant Objects.¹² The KAPTUR project¹³ investigated repository requirements for research data management in the visual arts and the Recollect plug-in¹⁴ transforms EPrints into a research data repository with an appropriate metadata profile. The Journal for Artistic Research¹⁵ (underpinned by the Research Catalogue database¹⁶) is another model of how to publish this research.

EPrints' software does not allow for sufficient modularization of the system to enable text and non-text outputs to be managed within the same repository. All plug-ins must agree a base metadata profile for the application – and while this could, using the Kultivate plug-in, modify the repository for better handling of arts research, it comes at the expense of text-based output types. The architecture of the Haplo repository enables different templates per output type with management of both non-text outputs and datasets in the same repository.

This functionality enabled us to introduce two refreshed non-text output item templates – the main one (Figure 4) used for output types except for exhibitions, which has additional fields (Figure 5), informed by work done by KULTUR,¹⁷ KAPTUR,¹⁸ REF metadata requirements and discussions with researchers, that reflect the form (rather than the format). Vocabulary was made more user-friendly – using creator and description, adding commissioning body to the publisher field, adding media type as a subfield and the introduction of a collaborator field recognizing other contributors. We embedded subcategories into text-based templates, for example 'exhibition catalogue' as an option within the book template.

'The architecture of the Haplo repository enables different templates per output type'

+ Creators

?

+ Collaborators

?

+ Description

+ Date

+ Keywords

?

+ Publisher or commissioning body

+ Files

Please note file size limit for upload using most browsers is 2GB

↑ Drag files here or [choose file...](#)

IMAGE CREDIT

MEASUREMENTS

RIGHTS

MEDIA TYPE

LICE

FILE ACCESS LEVEL

Closed (closed metadata and files)

Controlled (open metadata, closed files)

Open (open metadata and files)

Figure 4. Item template (used for artefact, composition, design, digital or visual media, performance outputs)

+ Title of work

+ Creators

?

+ Collaborators

?

+ Description

+ Date

+ Exhibition title

+ Keywords

?

+ Publisher or commissioning body

+ Files

Please note file size limit for upload using most browsers is 2GB

↑ Drag files here or [choose file...](#)

Figure 5. Exhibition template

7 Alongside these we introduced the portfolio template. This used the linked data model to allow existing output records to be connected into an overarching collection record to represent one larger piece of research. The initial template was further developed, and an enhanced portfolio template was added (Figure 6), enabling us to capture and make public practice research submitted to REF2021.

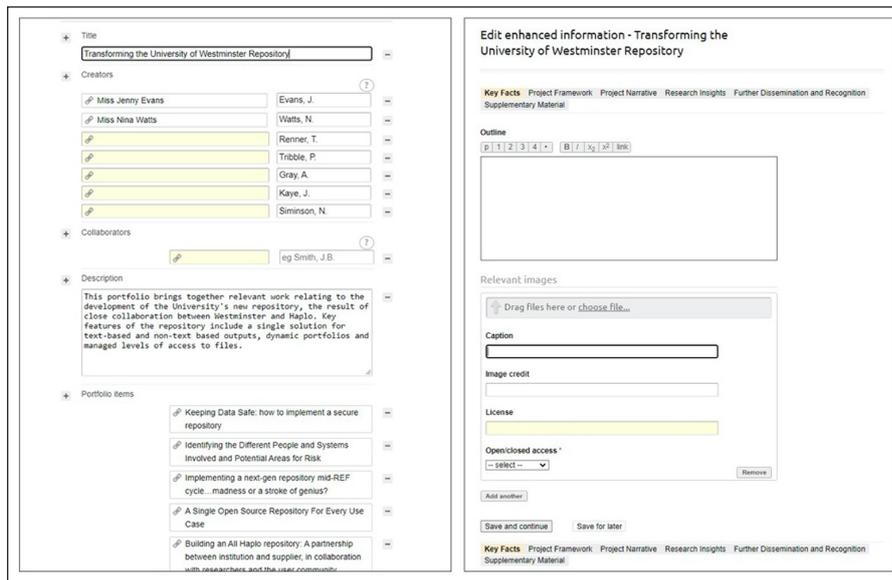


Figure 6. Portfolio template and enhanced portfolio template

The Haplo repository interface allows outputs to be displayed in a flexible manner, as mentioned above. This enabled us to give our whole repository a rebrand and practice research in the arts portfolios a clear and well-polished public presence (Figure 7). The portfolio records showcase not only the metadata and items within the portfolio, but also the research process leading to its creation. These outputs display alongside text-based outputs in a complementary format.

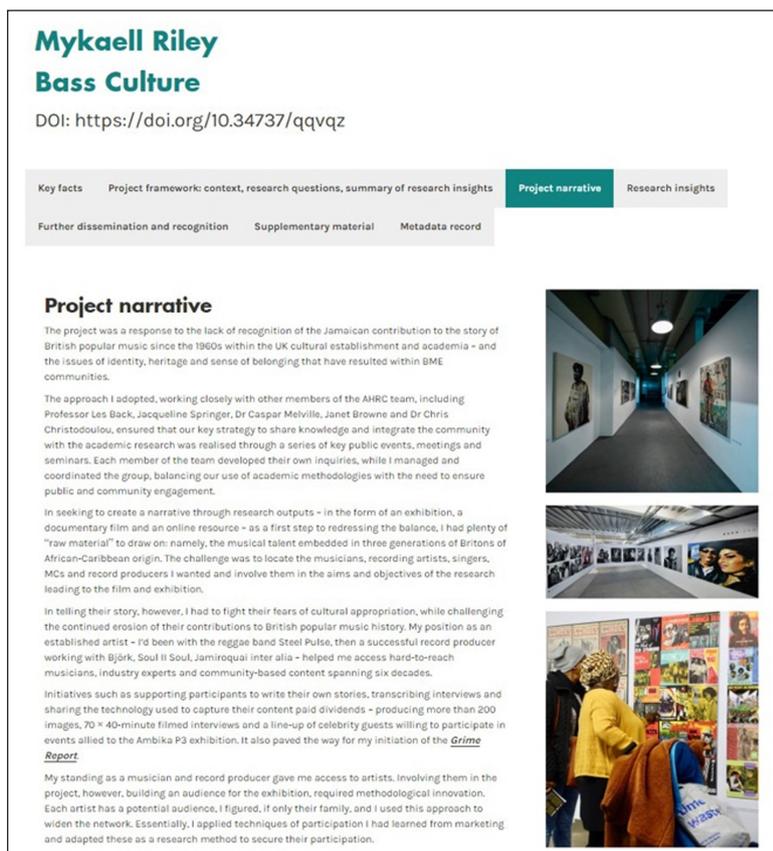


Figure 7. Public facing portfolio interface¹⁹

Datasets

The hybrid repository could not capture and provide access to research datasets since the public EPrints repository was unable to display them alongside text-based outputs. The hybrid approach meant that we needed the internal repository and workflow in Haplo to publish the data as approved internally to the public interface in EPrints. As a result of the EPrints system's inability to provide access to the datasets, the internal Haplo repository had to artificially replicate this limitation to prevent a situation where an item was approved for deposit but could not be deposited successfully. We needed the repository to capture datasets successfully and built in further functionality to provide a managed access workflow, secure storage and minting of DOIs.

Developing inclusive policy

External funder policy expectations continue to focus on traditional text-based outputs and datasets where grant funding enables the recovery of costs.²⁰ The UK Research and Innovation's (UKRI) new open access policy does now have a data access statement requirement in relation to research data, which includes a range of practice research outputs.²¹ We have made a conscious effort to develop policies that reflect the diversity of research outputs created by researchers at the University, with a focus on research activities and outputs rather than research papers or publications.

Making research open and FAIR in practice research disciplines is not as straightforward as in disciplines primarily publishing traditional text-based outputs. Sharing and making research 'open' is a combination of traditional open access where content can be licensed for reuse and data sharing where ownership of data varies. Intellectual property rights are more nuanced, with practice research outputs often not considered 'scholarly activity', defined in the University's Intellectual Property Policy²² as including the production of books, contributions to books, articles and conference papers.

'We have made a conscious effort to develop policies that reflect the diversity of research outputs created'

In 2019 the University's existing 'University Policy on Dissemination of Research and Scholarly Output' and open access mandate to deposit all doctoral theses in both WestminsterResearch and the UK's national thesis service, EThOS,²³ were updated to align with funder policies, to ensure the inclusion of all research outputs, and renamed the Open Access Policy.²⁴ The University's Research Data Management policy²⁵ (approved in June 2017 and under review in 2022), uses an inclusive definition of data, clearly referencing practice research and digital or physical objects and associated documentation. The acknowledgement that there are reasons why some research data cannot be shared is of relevance to practice research, which may have restrictions on sharing due to intellectual property rights.

Benefits

An increase in downloads! Figure 8 illustrates the increase over time in the total number of average downloads per month (split between the WestminsterResearch (EPrints) legacy repository and the WestminsterResearch (Haplo) repository) while recognizing that there were fewer items in the early years of the WestminsterResearch (EPrints) repository, which has contributed to the associated lower download figures.

The involvement of our practice researchers solidified a strong working relationship, which continued with the preparation of the University's submission for REF2021 and the release of our enhanced open access portfolios. It enabled us to develop a repository that comes much closer to representing their research than was previously possible.

Having one repository for all research outputs (while recognizing some research is better stored elsewhere) has enabled more holistic discussions around capturing and sharing research. Research outputs are more visible, can be reported on and include a more diverse range of research. This provides a much better foundation for future work in relation to responsible use of metrics and ensures a broader definition of research activities.

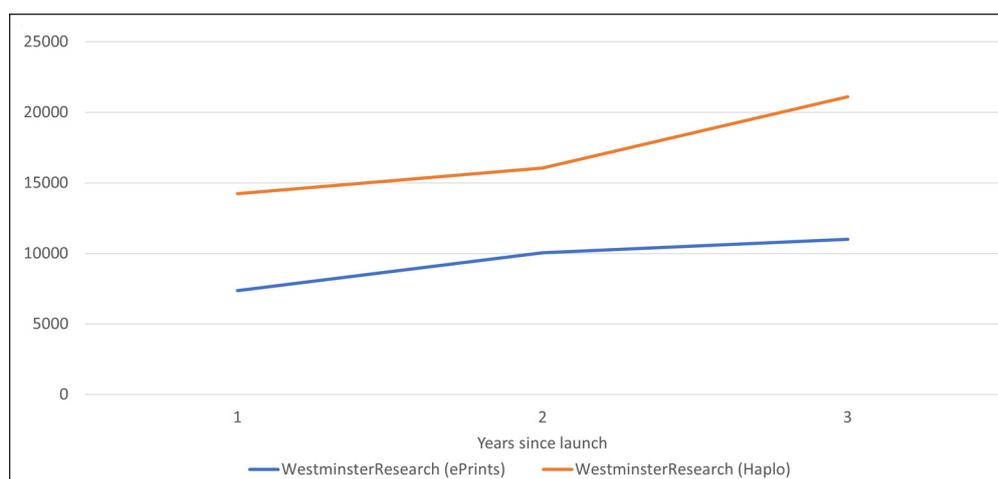


Figure 8. Average monthly downloads (3 Sept 2018 – 31 August 2021. Source: IRUS-UK)

The combination of flexible technology and having the wider research management system in place enabled the overlay of a REF2021 outputs module. This allowed relevant REF2021 metadata to be connected to repository records and made available to individuals with REF related roles.

It has also enabled the development of a machine-actionable Data Management Plans (DMPs) module based on the Research Data Alliance's common standard for machine-actionable Data Management Plans.²⁶ Adopting this in 2022–23 will highlight the benefits of using an integrated repository and research management system, tying together decisions made at the start of a research project with outputs deposited at the end of it.

An unexpected outcome has been the opportunity to engage with communities working on persistent identifier schema. This ecosystem exists to promote interoperability between systems, reduce manual data entry, save individual researchers' time and increase discoverability. It has, however, been developed with traditional publications in mind. It is much harder to make the case to researchers in practice research disciplines as they cannot 'see' the benefits for their research. This has led to collaboration with colleagues at Jisc, including discussions at international conferences,²⁷ and informed the questionnaire responses the team gave to the research underpinning the PRAG-UK Report 2 published in 2021.²⁸

Lessons learned

As a development partner there is such an opportunity to influence priorities for development. However, implementation can be challenging with legacy data to deal with, and some functionality going live after other clients, as happened with our ORCID integration and fixes to the workflow for datasets. Creative arts research does not neatly fall into structures and needs in-person follow up (even with user-friendly software) and we continue to work with researchers on a one-to-one basis.

Flexible software enabled functionality that researchers were not always ready for. The release of a REF2021 OA Policy compliance flag increased engagement about OA but confused many researchers who assumed their output could not be submitted to REF, rather than that there was no OA policy requirement for that output type. As a result, we hid this data from researchers but kept it available to those in research leadership roles to inform decision-making for the submission. We over-engineered a portfolio workflow that enabled editors to check and edit portfolios and this was eventually not used as the editing was done offline. A taxonomy for CREAM outputs was added, although a subsequent update based on their new digital strategy could not be implemented immediately. At the time of writing, this means the taxonomy is not used systematically as it does not reflect CREAM's current area of focus.

'Flexible software enabled functionality that researchers were not always ready for'

10 ‘Scope creep’ came in the form of the development of the REF outputs module. Having one single shared list of values for each linked field, for example funders, led to challenging discussions across the teams who support various elements of the wider product but also resulted in better connections and understanding between the different stakeholders.

While we commissioned and published a guide on copyright,²⁹ further conversations are needed in relation to licensing for reuse, with a more nuanced approach needed due to the intellectual property rights mentioned above.

At the time of writing, we have not addressed the systematic digital preservation of practice research outputs. Some practice research outputs are captured within our University Records and Archives digital preservation solution (hosted by Arkivum,³⁰ using open source software application Archivematica³¹) and can be accessed via the archive catalogue, Access To Memory³² (also open source). But this is often as a result of preservation being carried out to capture a representation of activity at a particular time, rather than focusing on preservation of the entirety of a research output.

Input from researchers guided the schema design and led towards a set of decisions and principles that have proven successful in other projects. The key lesson here was that to build trust and engagement with the community it is of critical importance to display fields using words familiar to the researcher. Metadata in repositories is historically focused on article-based research, which means that practice researchers are quickly discouraged when the system requests information in ways hard to understand, or with which their research ‘does not fit’. This was resolved by modifying field names between output types to be more contextually appropriate, for example ‘Publisher or commissioning body’, and internally mapping the relationship between these fields. The result of this was a system that presents users with the fields they expect, simplifying the deposit process while allowing the back end to translate to standard schemas – maximizing machine interoperability and providing a coherent view into the data for reporting purposes.

Conclusion

We focused on building an open source repository, that is standards based and meets the FAIR principles, that is also based on user needs, maintaining and improving support for more traditional text-based research, while introducing much needed equivalents for practice research in the arts and architecture. The flexibility of the architecture of the Haplo repository software has enabled us to meet multiple use cases in one place and save time for researchers. User engagement has been key, we continue to work in collaboration with our expert research community, with leadership and oversight by the University’s Research and Scholarly Communications team and in partnership with the development team at Haplo. This partnership has brought benefits not only to the University of Westminster but to the sector. The UK’s Arts and Humanities Research Council funded a scoping project in January 2022 to enable the review of this foundation work (both technology and standards) and the challenges it has raised, to identify how it could be expanded to other approaches to the capture of creative arts and practice research across disciplines, highlighting the intersectionality of practice research.

‘The flexibility ... of the ... repository software has enabled us to save time for researchers’

Note

1. Haplo Services Ltd became a division of Cayuse LLC in January 2021, which has seen the Haplo open source repository rebranded as the Cayuse open source repository. This article refers to Haplo for historical accuracy as it reflects work (mostly) carried out prior to the merger.

Abbreviations and Acronyms

A list of the abbreviations and acronyms used in this and other Insights articles can be accessed here – click on the URL below and then select the ‘full list of industry A&As’ link: <http://www.uksg.org/publications#aa>.

Competing interests

The authors have declared no competing interests.

References

1. "RCUK Common Principles on Data Policy," *Research Councils UK*, 2011, <https://web.archive.nationalarchives.gov.uk/ukgwa/20110518091755/http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx> (accessed 3 August 2022).
2. Ken Chad and Suzanne Enright, "The Research Cycle and Research Data Management (RDM): Innovating Approaches at the University of Westminster," *Insights* 27, no. 2 (7 July 2014): 147–53, DOI: <https://doi.org/10.1629/2048-7754.152> (accessed 3 August 2022).
3. 'EPrints Software – EPrints Services,' eprints, <https://www.eprints.org/uk/index.php/eprints-software/> (accessed 3 August 2022).
4. "Research Excellence Framework (REF) 2021 Open Access (OA) Policy," *REF2021*, 2019, https://www.ref.ac.uk/media/1228/open_access_summary_v1_0.pdf (accessed 3 August 2022).
5. Tim Berners-Lee, "Linked Data – Design Issues," *Linked Data*, 2006, <https://www.w3.org/DesignIssues/LinkedData.html> (accessed 3 August 2022).
6. Mark D. Wilkinson et al., "The FAIR Guiding Principles for Scientific Data Management and Stewardship," *Scientific Data* 3, no. 1 (15 March 2016): 160018, DOI: <https://doi.org/10.1038/sdata.2016.18> (accessed 3 August 2022).
7. "Open Archives Initiative Protocol for Metadata Harvesting," *Open Archives*, <https://www.openarchives.org/pmh/> (accessed 3 August 2022).
8. "ISO – ISO/IEC 27001 – Information Security Management," ISO, <https://www.iso.org/isoiec-27001-information-security.html> (accessed 3 August 2022).
9. Stephanie Meece, Amy Robinson, and Marie Therese Gramstadt, "Engaging Researchers With the World's First Scholarly Arts Repositories: Ten Years After the UK's Kultur Project," *New Review of Academic Librarianship* 23, no. 2–3 (3 July 2017): 209–32, DOI: <https://doi.org/10.1080/13614533.2017.1320767> (accessed 7 July 2022).
10. Victoria Sheppard and Wendy White, *KULTUR: Final Report*, (Jisc, 2009), <https://web.archive.southampton.ac.uk/kultur.eprints.org/Project%20Final%20report%20Mar%2009.pdf> (accessed 7 July 2022).
11. Marie-Therese Gramstadt, "Kultivating Kultur: Increasing Arts Research Deposit," *Ariadne: Web Magazine for Information Professionals*, 2012, <http://www.ariadne.ac.uk/issue68/gramstadt> (accessed 7 July 2022).
12. Tahani Nadim and Rebecca Randall, *Defiant Objects Project Report* (London: Goldsmiths, University of London, April 2013), <https://research.gold.ac.uk/id/eprint/8731/> (accessed 7 July 2022).
13. Marie-Therese Gramstadt et al., "KAPTUR the Highlights: Exploring Research Data Management in the Visual Arts," *Ariadne: Web Magazine for Information Professionals*, no. 71, 2013, <http://www.ariadne.ac.uk/issue71/garrett-et-al/#13> (accessed 7 July 2022).
14. Louise Corti et al., "ReCollect," *EPM*, 20 August 2014, <http://bazaar.eprints.org/367/> (accessed 8 July 2022).
15. "Journal for Artistic Research," *Journal for Artistic Research (JAR)*, <https://www.jar-online.net/en> (accessed 3 August 2022).
16. "Research Catalogue – an International Database for Artistic Research," *Society for Artistic Research*, <https://www.researchcatalogue.net/> (accessed 3 August 2022).
17. Victoria Sheppard, *KULTUR Project: Metadata Report* (Jisc, 2009), <https://web.archive.southampton.ac.uk/kultur.eprints.org/Metadata%20report%20Final.pdf>. (accessed 3 August 2022).
18. Leigh Garrett, Carlos Silva, and Marie-Therese Gramstadt, *KAPTUR: Technical Analysis Report* (Monograph: VADS Visual Arts Data Service: a Research Centre of the University for the Creative Arts, May 2012), <https://research.uca.ac.uk/1239/> (accessed 8 July 2022).
19. Mykaell Riley, "Bass Culture," Portfolio, University of Westminster, 2014, DOI: <https://doi.org/10.34737/qqvqz> (accessed 3 August 2022).
20. Erzsébet Tóth-Czifra, Laurent Romary, and Jennifer Edmond, "Towards a Plan(HS): DARIAH's Position on PlanS," 25 October 2018, DOI: <https://doi.org/10.5281/zenodo.3364398> (accessed 8 July 2022).
21. "UKRI Open Access Policy," UKRI, <https://www.ukri.org/publications/ukri-open-access-policy/> (accessed 8 July 2022).
22. "Intellectual Property Rights," University of Westminster, <https://www.westminster.ac.uk/about-us/our-university/corporate-information/policies-and-documents-a-z/intellectual-property-rights> (accessed 3 August 2022).
23. "British Library EThoS – Search and Order Theses Online," EThoS, <https://ethos.bl.uk/Home.do> (accessed 3 August 2022).
24. Nina Watts, "University of Westminster Open Access Policy," University of Westminster, 2018, DOI: <https://doi.org/10.34737/vx8w5> (accessed 3 August 2022).
25. Jenny Evans, "University of Westminster Research Data Management Policy," University of Westminster, 2017, DOI: <https://doi.org/10.34737/vx8w4> (accessed 3 August 2022).
26. Tomasz Miksa, Paul Walk, and Peter Neish, "RDA DMP Common Standard for Machine-Actionable Data Management Plans," *Zenodo*, 16 September 2020, DOI: <https://doi.org/10.15497/rda00039> (accessed 3 August 2022).

27. Jenny Evans, Taylor Mudd, and Adam Vials Moore, "Unheard Voices: Practice Based Arts Research and the PID Landscape," University of Westminster, 2021, DOI: <https://doi.org/10.34737/v5538> (accessed 3 August 2022); Adam Vials Moore et al., "Practice made Practical. Reducing the OTHER of non-text," Zenodo, 2021, DOI: <https://doi.org/10.5281/zenodo.5767094> (accessed 3 August 2022).
28. James Bulley and Özden Şahin, *Practice Research – Report 1: What Is Practice Research? And Report 2: How Can Practice Research Be Shared?* (London: Practice Research Advisory Group UK, 2021), DOI: <https://doi.org/10.23636/1347> (accessed 3 August 2022).
29. Jane Secker, "LibGuides: Copyright for Researchers: Introduction," University of Westminster, <https://libguides.westminster.ac.uk/copyrightresearchers/introduction> (accessed 3 August 2022).
30. "Data Archiving and Digital Preservation Solutions," Arkivum, <https://arkivum.com/> (accessed 3 August 2022).
31. "Archivematica: Open-Source Digital Preservation System," Archivematic, <https://www.archivematica.org/en/> (accessed 3 August 2022).
32. "AtoM: Open Source Archival Description Software," <https://www.accesstomemory.org/en/> (accessed 3 August 2022).

Article copyright: © 2022 Jenny Evans, Nina Watts, Taylor Mudd and Tom Renner. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use and distribution provided the original author and source are credited.



Corresponding author:

Jenny Evans
 Research Environment and
 Scholarly Communications Lead
 University of Westminster, GB
 E-mail: J.Evans2@westminster.ac.uk
 ORCID ID: 0000-0001-9455-3658

Co-authors:

Nina Watts
 ORCID ID: 0000-0002-3773-6797

Taylor Mudd
 ORCID ID: 0000-0002-9980-613X

Tom Renner
 ORCID ID: 0000-0001-9500-4407

To cite this article:

Evans J, Watts N, Mudd T and Renner T "From legacy to next generation: a story of collaboration to push the boundaries of the Haplo (now Cayuse) open source repository software," *Insights*, 2022, 35: 14, 1–12; DOI: <https://doi.org/10.1629/uksg.582>

Submitted on 22 April 2022

Accepted on 09 June 2022

Published on 07 September 2022

Published by UKSG in association with Ubiquity Press.