Mapping the Landscape of Digital Research Tools: Recommendation Package

Working Group name: RDA-OfR Mapping the Landscape of Digital Research Tools

Working Group co-chairs: Emmanuel Adamolekun (0000-0003-2992-5448), Francis P. Crawley (0000-0002-6893-5916), Rory Macneil (0000-0002-8429-096X), Adam Vials Moore (0000-0002-2085-1908) and Hea Lim Rhee (0000-0002-4171-5710)

Assisting co-chairs: Marcelo Garcia 0000-0002-2927-2371 & Richard Pitts 0000-0002-2037-3360

Working Group Facilitator and Editor: Connie Clare (0000-0002-4369-196X)

Version: July 2024 | CC-BY 4.0 International |

Table of Contents

Introduction and Impact

The digital research data infrastructure landscape comprises a myriad of tools for managing and sharing research data during various stages of the research data lifecycle (RDL). Such research tools vary widely depending on data type, user requirement, provider, and subject area. In the context of the Mapping the Landscape of Digital Research Tools Working Group (WG), research tools enable researchers to perform one or more operations, typically on data, and often with data as the output. Tools are usually intended for use by humans. In this context we are explicitly excluding physical instruments.

The diversity and variety of research tools can prove overwhelming and challenging for stakeholders working within the digital research data ecosystem to understand, navigate, and select the most appropriate tool to meet their needs and objectives. The categorisation of research tools, based on their features, functionalities and how they interoperate, remains unclear. In many cases, research tools are not interoperable, often leading to siloed working within organisations and disciplines, thereby limiting the scope of research and the ability to share and reuse data.

This WG aimed to address these challenges by: (i) categorising different types of research tools; and, (ii) mapping different types of research tools to the RDL based on their features and functionalities. The WG produced a categorisation schema (a conceptual framework) of research tool types that includes terminologies, definitions, and associated metadata describing features and functionalities of different tool types. In order to ensure the categorisation schema of tool types is accessible, engaging, and interactive, the WG has produced a visual online map of the digital research tool landscape which serves as a dynamic prototype to be further developed in the future.

Adoption Use Cases

To our knowledge, this RDA WG is the first initiative of its kind to categorise different types of research tools with a primary focus on their utility within the RDL. Providing the global research data community with an interactive map of the digital research data tool landscape that can be navigated according to the RDL represents a novel approach to characterising the research data ecosystem. The deliverables produced by this WG aim to provide value and impact for the following adopters:

Adopter	Value/Impact
Researchers (e.g., data creators and users)	To understand, navigate, and select suitable research tools for managing and sharing data by providing information about their functionalities, relevance, and applicability to the various stages of the RDL.
Data support professionals (e.g., data managers)	To gain improved understanding of the digital research data infrastructure landscape, and become better equipped with essential knowledge of different types of research tools to provide relevant support, training, and education.
Open Science/Research/Data Commons professionals	To understand the features, functionalities, and interoperability of different types of research tools that can be used within diverse marketplaces or ‘commons’ for data and services.
Tool developers/ providers	To: (i) understand the different research tools operating within the digital research data landscape; and, (ii) improve tool features, functionalities, harmonisation, and interoperability to enhance data management and sharing practice.
Research performing organisations	To make informed recommendations at the organisational policy level to staff regarding appropriate types of research tools for the management and sharing of research data.
Publishers	To make informed recommendations to authors and journal editors regarding appropriate types of research tools for the

	management, publication, and sharing of data associated with journal manuscripts.
Funders	To make informed recommendations to researchers and project managers based on data management plans for funded research.

Potential Adopters

- The California Digital Library and Research Space are recipients of a US National Science Foundation grant to hold a workshop on the topic of 'Vertical Tool Interoperability. The MaLDReth model is included as a core structuring element for the workshop, around which issues relating to interoperabilty between tools used in different stages of the research data lifecycle can be discussed. The workshop will take place in late 2024 or early 2025..
- JISC, the UK digital, data and technology agency focused on tertiary education, research and innovation, is interested in using the MaLDReTH model as a reference for policies relating to recommendations and support for research tools and services.
- Discussion with the GORC II WG is scheduled to explore synergies and possible integration between the MaLDReth model at the Services and Tools Element of the GORC model. The first step in this process will take place in September/October 2024, with an initial exploration of mapping between the two models.

Deliverable 1: The creation of a research data lifecycle (RDL) model and crosswalk to existing models

Executive Summary

As there is not one accepted universal Research Data Lifecycle (RDL) model, the first deliverable by the RDA-OfR Mapping the Landscape of Digital Research Tools Working Group (WG) involved the creation of a harmonised RDL that could be used as the foundational framework for underpinning later work on the categorisation and characterisation of digital research tools.

To achieve this, a landscape review was carried out to identify a list of existing RDLs. A qualitative analysis process was conducted to determine the top five best characterised and most comprehensive RDLs. Their stages and definitions were aligned using a semantic distance methodology over an etymological/taxonomic net to produce the Mapping the Landscape of Digital Research Tools Harmonised (MaLDReTH) RDL model.

Documentation

D1_The creation of a harmonised research data lifecycle (RDL) model and crosswalk to e…

Deliverable 2: The identification, categorisation, and mapping of different types of research tools: A categorisation schema

Executive Summary

The second deliverable by the RDA-OfR Mapping the Landscape of Digital Research Tools Working Group (WG) involved the creation of a categorisation schema of digital research tools.

The first phase involved a comprehensive landscape review, conducted over five months (from September 2023 to January 2024), whereby several task groups convened during virtual meetings, plus an extended in person working meeting during the RDA's 21st Plenary in Salzburg, to collate a list of representative digital research tools used during each of the 12 stages of the MaLDReTH RDL model. The number of representative tools listed for each RDL stage varied in length, from four to five tools in the shortest list and 40+ tools in the longest.

The second phase, carried out by a subgroup of four WG members from January to April 2024, involved validating the lists of tools and identifying common categories (types) of tools used during each RDL stage. Due to time constraints, a maximum of three tool categories were identified for each RDL stage. During the final step, the subgroup identified a maximum of three individual digital research tools that serve as representative examples for each tool category.

Documentation

D2_The identification, categorisation, and mapping of different types of research tools: A c…

Deliverable 3: The creation of a preliminary structural framework for an online open access 'map of the digital research tool landscape'

Executive Summary

The final deliverable of the RDA-OfR Mapping the Landscape of Digital Research Tools Working Group (WG) undertakes the necessary foundational work required to create a preliminary structural framework for an online open access 'map of the digital research tool landscape'. In essence, this 'map' is an open access autonomous relational database hosted by the RDA

Foundation (as a legal entity on behalf of the RDA), owned by the community, and powered by Oracle for Software. This arrangement has been discussed and agreed by the RDA and Oracle for Research.

The database, navigable by research data lifecycle stage, contains searchable information (e.g., features, functionalities, interoperability) about different types of digital research tools, and allows for ongoing community curation and further development.

Based on Deliverable 1 (The creation of a harmonised research data lifecycle (RDL) model and crosswalk to existing models) and Deliverable 2 (The identification, categorisation, and mapping of different types of research tools: A categorisation schema), this deliverable presents a prototype for the database in addition to recommendations for its long-term maintenance, sustainability, and adoption.

The idea is that this tool prototype will be further developed and implemented by a follow-on RDA WG, considering the recommendations detailed in this deliverable to ensure the final database remains up to date with newly emerging types of digital research tools and evolves with the ever-changing digital research data infrastructure landscape.

Documentation

D3_The creation of a preliminary structural framework for an online open access ‘map of t…

Tool Prototype:

https://adamvialsmoore-jisc.shinyapps.io/r-proto

Recommendations and Future Work

Recommendations

Based on the findings and feedback gathered during the development and community engagement processes, the following recommendations are proposed:

Adoption and Integration

1. Facilitate and encourage widespread adoption of the tool by promoting the tool’s value for research institutions, data management platforms, and relevant professional networks. Develop targeted outreach strategies to engage different stakeholder groups

and highlight the tool's benefits, such as live demonstrations, webinars, conference presentations and posters.

2. Integrate the tool with existing digital research infrastructures to enhance its accessibility and usability. In particular, discuss synergies and potential integration with the Services and Tools Element of the GORC Model. Collaborate with other research tool developers and platform providers to facilitate seamless integration and data interoperability.

3. Identify an appropriate hosting platform for the tool. Since hosting costs will be minimal, it should be possible to identify a viable solution. Discuss with the Secretariat the possibility of hosting on the RDA website.

Continuous Improvement

4. Establish a process for regularly updating the tool's database to include new and emerging digital research tools. Monitor trends and advancements in digital research tools to ensure the tool remains current and relevant.

5. Continuously solicit feedback from users throughout the development and implementation process to identify areas for improvement and potential new features for user-driven enhancement. Implement a feedback loop to incorporate user suggestions and enhancements in future iterations of the tool.

Expanding Functionality

6. Enhance the tool's search and filtering capabilities to provide users with more granular control over their searches. Implement advanced filtering options based on specific criteria such as tool functionality, research domain, and compatibility/interoperability with other tools.

7. Explore opportunities for integrating the tool with other research management and data sharing platforms. Develop API functionalities to enable seamless data exchange and interoperability with external systems.

Educational and Support Resources

8. Develop comprehensive user guides and tutorials to assist users in navigating and utilising the tool effectively. Provide step-by-step instructions and best practice recommendations to enhance user experience based on stakeholder value and interest.

9. Organise generic training sessions and workshops to educate users about the tool's features and functionalities. Provide hands-on, practical training opportunities to help users become proficient in using the tool for their individual research data management needs.

Future Directions

Several strategic initiatives and future directions are proposed to build on the success of Deliverable 3, and ensure the ongoing development and impact of the online tool.

Database Expansion

Continuously expand the tool's database to include different types of digital research tools with relevant examples. Collaborate with tool developers and research communities to identify and categorise new tools.

Global Outreach

Promote the tool globally and establish partnerships with key stakeholders. Engage with international research communities and organisations to promote the tool and encourage global adoption. Establish partnerships with key stakeholders to facilitate knowledge sharing and collaboration.

Sustainability Plan

Develop a sustainability plan to ensure the long-term maintenance and support of the tool. Explore funding opportunities and collaborative ventures to secure the necessary resources for ongoing development.

Note: There has been ongoing discussion within the WG about the choice of software underpinning the online tool. The prototype presented herein currently uses open source software on open storage/facilities. It's coded entirely in R / Shiny with non commercial libraries. The WG should consider the future production status of the tools developed, and investigate the advantages and disadvantages of using different software. For example, should the final application be open or closed? What about the schema for the information stored? Should the storage layer be in a modern, stable environment like Oracle OCI or remain as a portable SQlite file?

Innovative Features

Investigate the potential for incorporating advanced features such as machine learning algorithms to provide personalised tool recommendations. Explore the integration of

visualisation tools to enhance the user experience and provide insights into the research data lifecycle.

Next steps

Members of the WG and external collaborators are in the process of establishing a follow-on RDA WG to continue the development of the online tool. The focus of the follow-on WG will be to:

- **Continue development of the tool:** This will include the adding more information about the attributes of different types of digital research tools, such as whether they are open or closed source, disciplinary or agnostic, and, interoperable or not.

- **Explore sustainability of the model:** This will include ways in which it can be kept up to date to reflect changes in technology, development of new research tools, changes in research practices, etc. Also provide options for continuing support for the model, which could include participation by various RDA groups and external organisations and/or communities.

- **Collaboration with RDA groups working on the Global Open Research Commons (GORC):** As the tool prototype was developed in recent months, the Mapping the Landscape of Digital Research Tools WG had discussions with the Global Open Research Commons IG and GORC International Model WG about the synergies in work and potential for collaboration. Task Group 5 of the GORC International WG has undertaken an extensive literature review and released a Commons Attributes Model (Version 0.5) that identifies a suite of services and tools that will inform the work of this WG. Efforts to describe the features, functionality, and interoperability of different types of research tools will complement the development of the 'Commons Integration Roadmap' (GORC WG Deliverable) by providing key information about different types of research tools, and highlighting areas for the improvement of their interoperability and user experience.

Conclusion

Deliverable 3 has successfully developed a prototype for an interactive online tool that maps digital research tools to the research data lifecycle.
This tool provides significant value to researchers, data managers, tool developers, research organisations, funders, and publishers. The comprehensive feedback from the RDA community has been instrumental in refining the tool and ensuring its relevance and usability.

The recommendations and future directions outlined in this deliverable report offer a roadmap for continued development, adoption, and impact. By leveraging the work achieved by the current WG, as well as the insights and feedback gathered for the creation of this deliverable, a follow-on RDA WG is well-positioned to enhance the tool's functionality and expand its reach, ultimately contributing to the advancement of digital research data management and sharing practices.

For further details, please refer to the RDA-OfR Mapping the Landscape of Digital Research Tools WG.