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

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)

Table of Contents

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) and owned by the community. 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.

Aims and Objectives

The WG aimed to create the preliminary structure (a tool prototype) for an open access 'map of the digital research tool landscape' navigable by research data lifecycle stage. The ultimate aim was to provide the data community with a dynamic, user-friendly, online database that contains searchable information (e.g., features, functionalities, interoperability) about different types of research tools. The database should also permit ongoing community curation so that it remains up-to-date with newly emerging types of research tools and evolves with the ever-changing digital research data infrastructure landscape. This may include significant data and software-related developments, e.g., Artificial Intelligence (AI).

It was beyond the scope of the current WG to create the final version of the database, however, the objective of Deliverable 3 was to develop a prototype of an autonomous relational database. With the foundational work in place, the idea is that a follow-on RDA

WG will continue to develop and implement the database, considering the recommendations for its long-term maintenance, sustainability, and adoption presented herein.

It is acknowledged that many different 'maps' of the digital research infrastructure exist. This particular WG endeavoured to create an open access interactive map that facilitates an intuitive, engaging, and explorative experience of navigating the digital research tool landscape.

Methodology and Results

Tool Prototype Development

Design and Development
From April to July 2024, Adam Vials Moore (WG co-chair), led the prototype design and development process. This was an iterative process that began with outlining the prototype's information model, core functionalities and user interface requirements (Figures 1 to 5). The design prototype incorporated essential features, such as categorisation, search and filter options. A set of wireframes were developed (Figure 4) to show the general user journey and affordances/interrelation of application routes and actions.

Technical Development
Modern web technologies (initially R/Shiny then python/flask) were used to build the tool, ensuring that it is scalable, responsive and user-friendly. A robust database (Figure 1) was implemented to store information about the different digital research tools and their categorisations, based on Deliverable 2.

Testing and Iteration
Multiple iterations of implementation testing were conducted through demonstrations to identify and resolve technical issues or usability challenges.
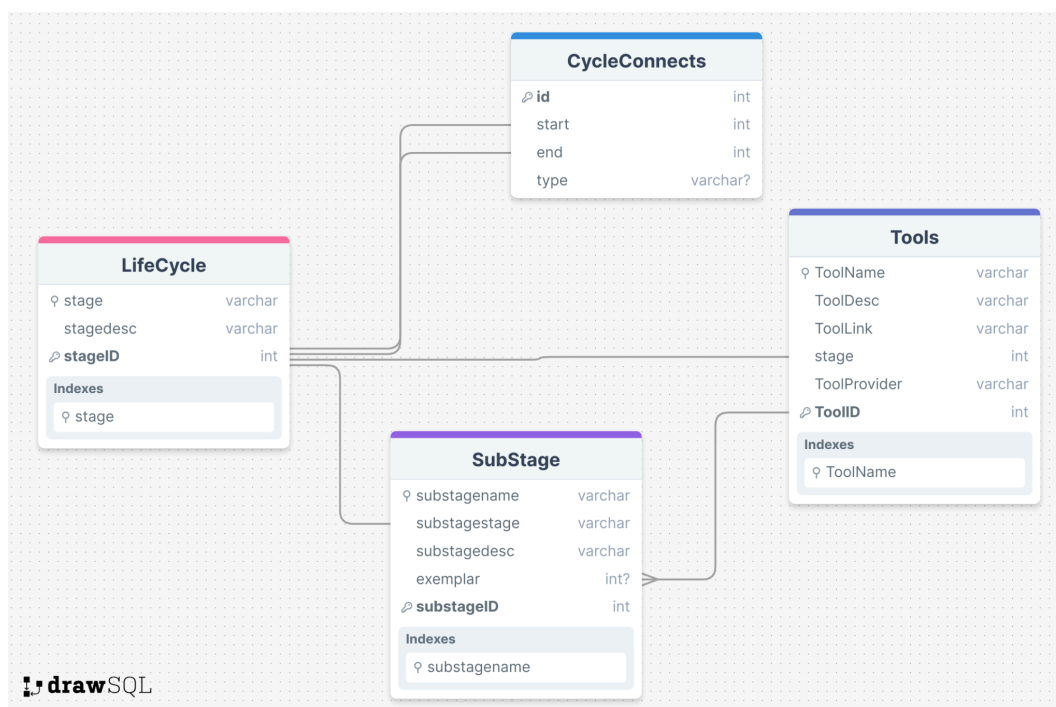
```sql
CREATE TABLE `LifeCycle`(
    `stage` VARCHAR(255) NOT NULL,
    `stagedesc` VARCHAR(255) NOT NULL,
    `stageID` INT NOT NULL,
    PRIMARY KEY(`stageID`)
);
ALTER TABLE
    `LifeCycle` ADD INDEX
`lifecycle_stage_index`(`stage`);
ALTER TABLE
    `LifeCycle` ADD INDEX
`lifecycle_stage_index`(`stage`);
CREATE TABLE `Tools`(
    `ToolName` VARCHAR(255) NOT NULL,
    `ToolDesc` VARCHAR(255) NOT NULL,
    `ToolLink` VARCHAR(255) NOT NULL,
    `stage` INT NOT NULL,
    `ToolProvider` VARCHAR(255) NOT NULL,
    `ToolID` INT NOT NULL,
    PRIMARY KEY(`ToolID`)
);
ALTER TABLE
    `Tools` ADD INDEX
`tools_toolname_index`(`ToolName`);
ALTER TABLE
    `Tools` ADD INDEX
`tools_toolname_index`(`ToolName`);
CREATE TABLE `SubStage`(
    `substagename` VARCHAR(255) NOT NULL,
    `substagestage` VARCHAR(255) NOT NULL,
    `substagedesc` VARCHAR(255) NOT NULL,
    `exemplar` INT NULL,
    `substageID` INT NOT NULL,
    PRIMARY KEY(`substageID`)
);
ALTER TABLE
    `SubStage` ADD INDEX
`substage_substagename_index`(`substagename`);
ALTER TABLE
    `SubStage` ADD INDEX
`substage_substagename_index`(`substagename`);
ALTER TABLE
    `SubStage` ADD CONSTRAINT
`substage_substagestage_foreign` FOREIGN
KEY(`substagestage`) REFERENCES
`LifeCycle`(`stageID`);
ALTER TABLE
    `Tools` ADD CONSTRAINT
`tools_stage_foreign` FOREIGN KEY(`stage`)
REFERENCES `LifeCycle`(`stageID`);
ALTER TABLE
    `SubStage` ADD CONSTRAINT
`substage_exemplar_foreign` FOREIGN
KEY(`exemplar`) REFERENCES `Tools`(`ToolID`);
```

Figure 1. Information Schema of Underpinning Information

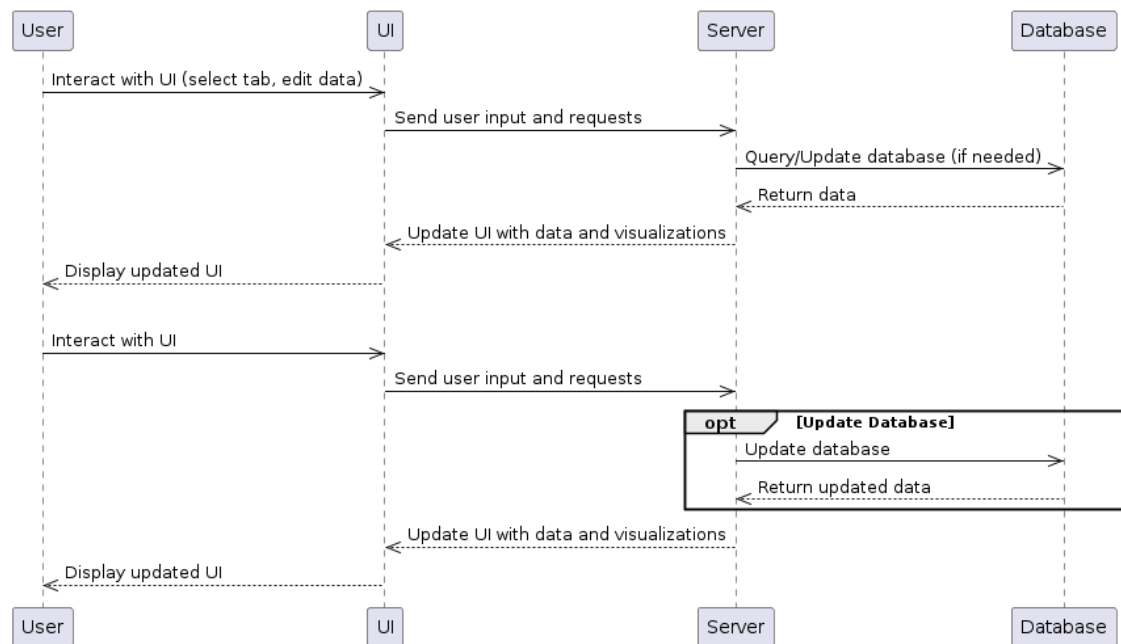**Figure 2.** Unified Modelling Language (UML) for Database User Journey

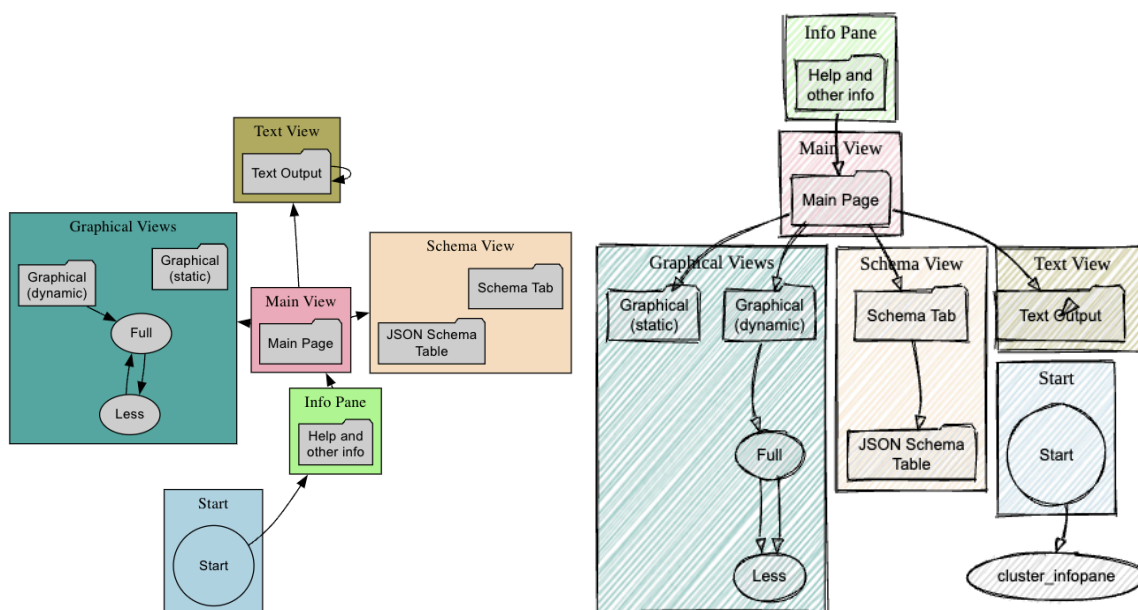(a)   (b)

Figure 3. (a) Graph view of prototype UI/UX journey and **(b)** Vertical affordance tree view of prototype interactions
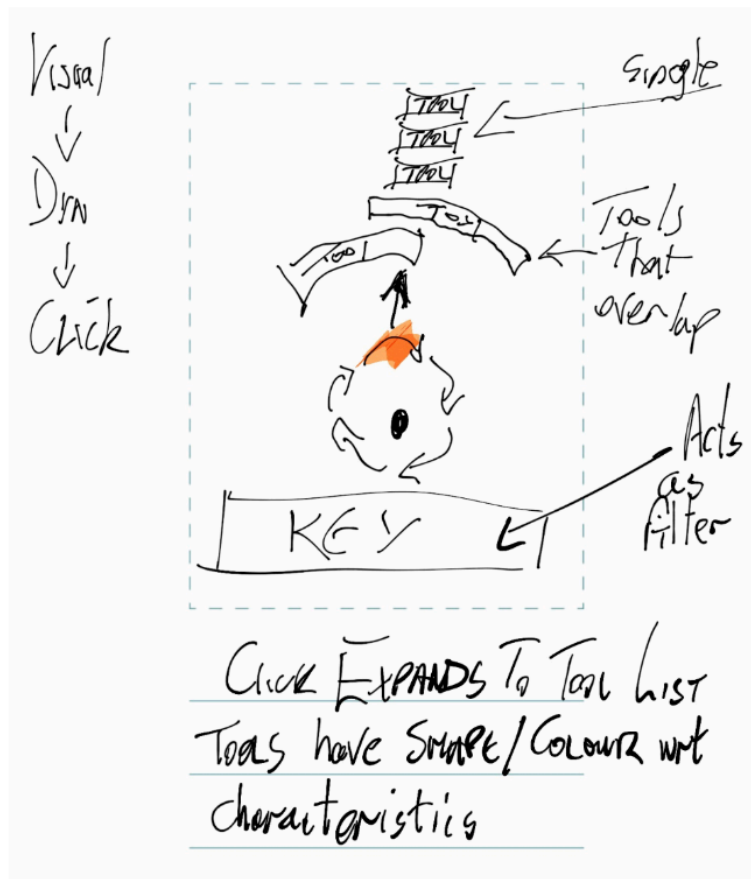
Wireframes

Figure 4. Example wireframe for the tool prototype.

See all wireframes: Data cycle tools vizn wireframes.pdf

Implementation (R/Shiny)

Following on from the work to define the features and functionality of the application, an initial prototype was built in R/Shiny. The prototype used graphViz interpretation to display the MaLDReTH lifecycle. It used a subset of the full RDL tool information to demonstrate display of the information both in text and schema view.

https://adamvialsmoore-jisc.shinyapps.io/r-proto/
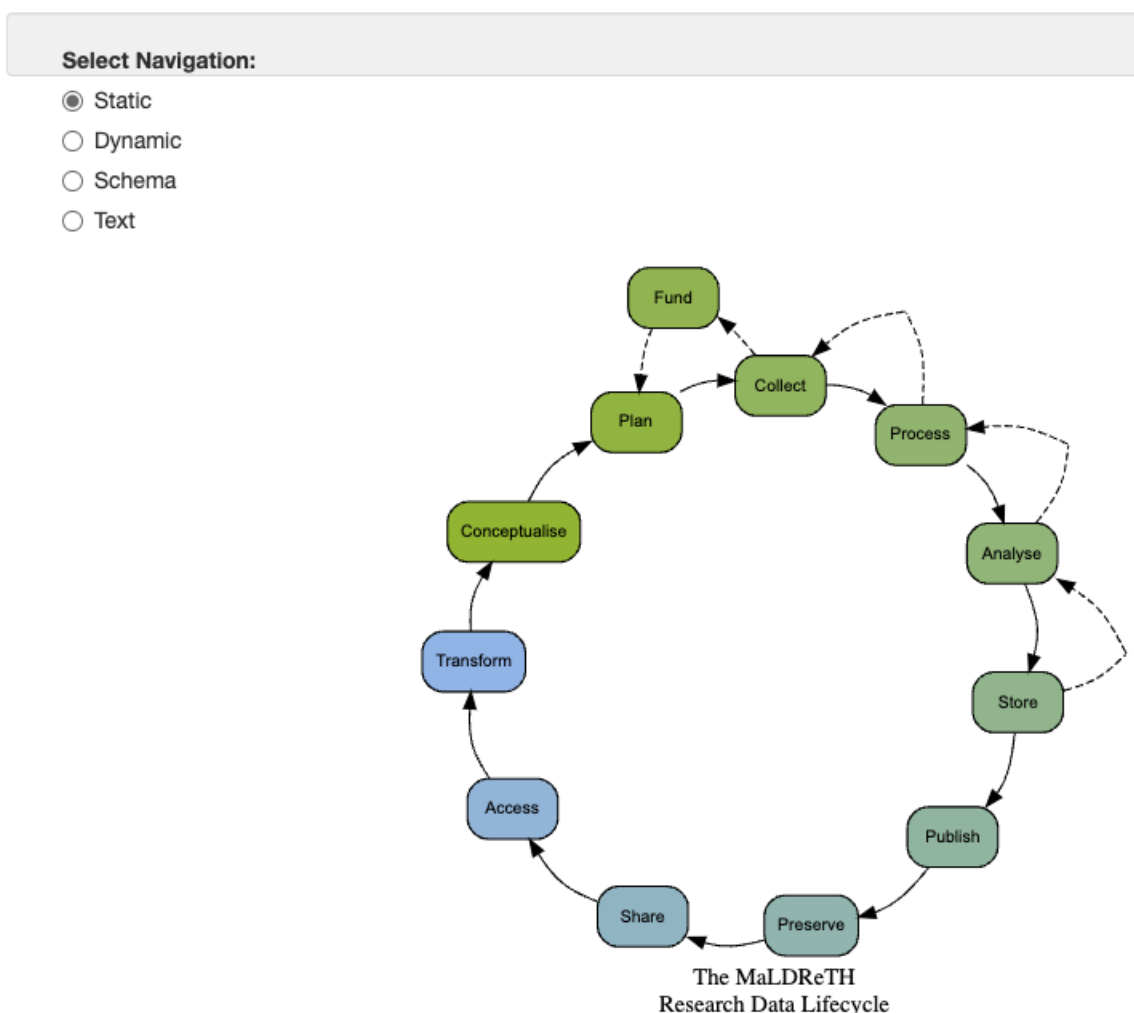
Figure 5. Open access tool prototype (Model 1), to map different types of digital research tool to the Mapping the Landscape of Digital Research Tools Harmonised (MaLDReTH) Research Data Lifecycle (RDL) model.

From there a User interface approaches have been developed which are displayed below

Model 2 (overview **Figure 6)** is backed by a SQLite database and presents our vision of how data can be maintained from this approach, but the User interface lacks polish.
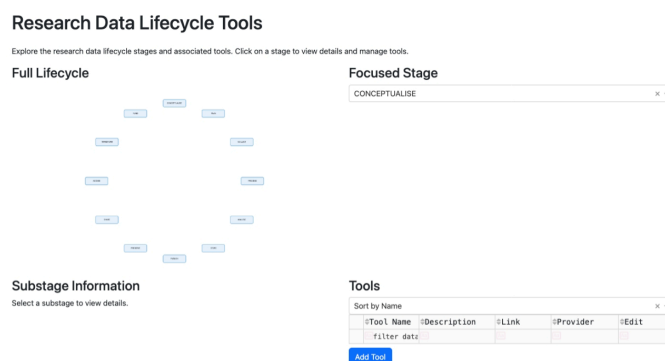
Figure 6

Model 2 - **Figure 7** interacting with the full life cycle shows the information related in the other windows

Figure 7

By pressing the add tool one can then interact with the SQLite database **Figure 8**. We see this sort of easy editing and data maintenance very important for the future longevity of the project.

Figure 8

Model 3 **Figure 9** and Model 4 **Figure 10** below show our further user interface experiments. This is ongoing project work which we fully expect to continue into the next working group in this area.

Research Data Lifecycle

Click on a stage to view its substages. Click on a substage to view its exemplars. Hover over elements to see descriptions. The connections between stages show the flow of the research lifecycle.
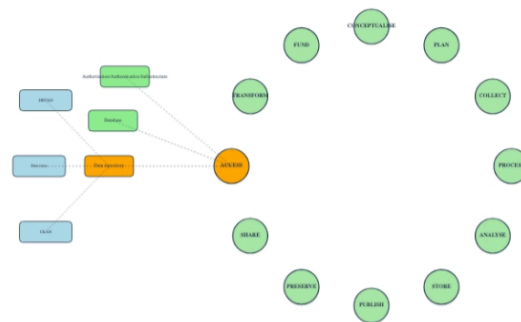
How to use this visualization:

1. The main circle shows the stages of the research data lifecycle.
2. Dotted lines with arrows show the connections and flow between stages.
3. Click on any stage (green circles) to view its substages.
4. Click on a substage (light green rectangles) to view its exemplars.
5. Use the zoom controls in the top right corner to adjust the view.
6. Hover over any element to see its description.
7. Click on a selected stage or substage to deselect it.
8. The orange highlight indicates the currently selected item.

Figure 9 - Model 3 Interface trial

Figure 10 Model 4 - Another approach to the User Interface.

As stated above this is a work in progress. Comments on the various models are welcome!

Community Engagement

Members of the RDA community were engaged at all stages of the tool prototype development process. Presentations were delivered and feedback was solicited from diverse stakeholders (researchers, data support professionals, software developers and funders) during the RDA's 20th Plenary in Gothenburg, 21st Plenary in Salzburg and 22nd virtual Plenary, in addition to the monthly WG meetings. Collaborative notes, meeting recordings and presentation slides are documented on the WG landing page.

During such meetings, live demonstrations of the tool prototype features and functionalities were conducted, and feedback was solicited via open discussions. Feedback was collated and analysed by a WG subgroup to identify challenges and opportunities for improvement (see 'Key Findings'). The categorisation schema of digital research tools (Deliverable 2) was also validated through extensive user feedback, confirming its accuracy and relevance as the foundation for the tool prototype.

Reporting and Recommendations

This document consolidates feedback from community engagement and formulates recommendations for future work and adoption strategies.

Key Findings

The development and community engagement processes yielded several important findings that highlight the value and potential impact of the online tool.

Tool Effectiveness

The prototype presents a foundation for an online tool that effectively maps digital research tools according to their features, functionalities, and relevance to the research data lifecycle. Based on feedback following live demonstrations of the tool prototype, users found its principles and approach compelling. It was described as intuitive and easy to navigate, enabling users to efficiently identify and explore different types of digital research tools for specific stages of the research data lifecycle.

Value to Stakeholders

Based on user feedback, the tool was deemed to provide value to the following stakeholders:

Researchers - *Enhanced ability to manage and share data.*
The tool provides researchers with a comprehensive resource for discovering and selecting suitable research tools, enhancing their ability to manage and share data effectively.

Data Managers - *Improved understanding of the digital research data infrastructure.*
Data support professionals benefit from an improved understanding of the digital research data infrastructure landscape, enabling them to provide more relevant and targeted support.

Tool Developers - *Insights into tool features and interoperability.*
Developers gain insights into the features, functionalities, and interoperability of different research tools, allowing them to improve their offerings.

Research Performing Organisations - *Informed policy recommendations.*
Organisations can make informed policy recommendations regarding the use of specific research tools for data management and sharing.

Funders and Publishers - *Better guidance for data management practices.*
Funders and publishers can provide better guidance to researchers and authors, ensuring that data management practices align with best practices and standards.

Feedback for Improvement

Users provided feedback on potential enhancements, such as additional filtering options, integration with other research platforms, and the inclusion of more tools in the database. It was also suggested that a follow-on RDA WG should research and explore other existing work undertaken to map the digital research data infrastructure landscape, and to harmonise efforts if suitable. User feedback also highlighted the importance of continuous updates to ensure the tool remains relevant and comprehensive.

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](#).

Contributors

	Name	Affiliation	ORCID
1	Adam Vials Moore	Jisc	[0000-0002-2085-1908](#)
2	Marcelo Garcia	King Abdullah University of Science and Technology (KAUST)	[0000-0002-2927-2371](#)

3	Rory Macneil	Research Space	0000-0002-8429-096X
4	Richard Pitts	Oracle for Research	0000-0002-2037-3360