



Creating an Implementation Plan for Harvestable Metadata Services on Twine

A USER's GUIDE

Winnie Ak Wai Li

Research Associate for the International Technology Office

winnieawli@oceannetworks.ca

Chantelle Verhey

Research Associate for the International Technology Office

winnieawli@oceannetworks.ca

Dr. Alicia Urquidi Díaz

Research Associate for the International Technology Office

aurquidi@oceannetworks.ca

Dr. Karen Payne

Associate Director for International Technology

kpayne@oceannetworks.ca

A REPORT BY THE

World Data System • International Technology Office

Funded under the Canadian New Digital Research Infrastructure Organization (NDRIO)

collaborative agreement number 51387-50326

Cite this report:

Li, W. A. W., Verhey, C., Urquidi Díaz, A., & Payne, K (2021). Creating an implementation plan for harvestable metadata services on Twine: A user's guide. World Data System - International Technology Office. <https://doi.org/10.5281/zenodo.4589036>



This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.





Ocean Networks Canada
University of Victoria
#100-2474 Arbutus Rd. Victoria, BC V8N 1V8
P 250 472 4527 | wds-ito.org

Table of Contents

Executive Summary	4
Background	5
Introduction	6
The Interactive Tool	7
The User Mode	7
The Developer Mode	15
Limitations and Future Development	20

Executive Summary

The World Data System's Creating an Implementation Plan for Harvestable Metadata Services on Twine: A user's guide is made up of this guidance document and an associated [interactive tool](#). This guidance document is designed to instruct users how to use the associated interactive tool, while the interactive tool is designed to aid research data repository managers in creating a semi-customized implementation plan for harvestable metadata services based on the current state of their research data repositories.

This guide covers the background of this project, the introduction of this guide, recommendations for both data managers and developers on how to work with the interactive tool, and the limitations and future development. For users whose repositories have developed harvestable metadata services, this resource may help them to expand their services to dataset metadata aggregators like Federated Research Data Repository (FRDR)¹, and/or web crawlers such as Google Dataset Search² to increase the findability of their repositories. For users whose repositories haven't developed harvestable metadata services, this resource may also help them to find, assess and select the metadata standards and exposure protocols that would make datasets interoperable within the infrastructures that are currently used in the dataset metadata aggregators and/or web crawlers.

¹ Federated Research Data Repository, "Federated Research Data Repository," 2021, <https://www.frdr-dfdr.ca/repo/>.

² Google, "Google Dataset Search," <https://datasetsearch.research.google.com/>.

Background

This guide is an outcome of the Harvestable Metadata Services Working Group (HMetS-WG) of the World Data System-International Technology Office (WDS-ITO), which ran between October 2019 and March 2021. It was created in response to the interests expressed in the World Data System Member Survey 2019³ as a community-driven, multidisciplinary platform where members of the WDS could interact, collaborate and share their good practices, approaches and pain points associated with creating harvestable metadata services.

The HMetS-WG was supported by the Canadian funding agency New Digital Research Infrastructure Organization (NDRIO), in a collaborative project with the WDS-ITO and Ocean Networks Canada (ONC) to connect Canadian digital research infrastructure (DRI) with global initiatives and practices.

Adherence to international standards and best practices is a key tenet for successful DRI integration and interoperability. As an established and widely-used tool to support effective data discovery and data sharing, harvestable metadata is often the first step repositories take in their maturation from locally held data and metadata. Further, harvestable metadata is an indicator of dataset findability⁴ and, as it entails better integration with the wider data management community, it is a measure of repository TRUST worthiness.⁵

This guide is one of a suite of products created as part of this NDRIO-funded project, which includes other guidance documents targeted to data repository managers, including a [flow chart](#) to help create a plan to provision harvestable services, a [template implementation plan](#),⁶ and a [curated collection of reference materials](#) that cover issues associated with harvestable

³ Karen Payne and Alicia Urquidí Díaz, “World Data System Member Survey 2019” (Victoria, BC: Zenodo, June 8, 2020), <https://doi.org/10.5281/ZENODO.3840406>.

⁴ CoreTrustSeal Standards And Certification Board, “CoreTrustSeal Trustworthy Data Repositories Requirements 2020–2022” (Zenodo, 2019), <https://doi.org/10.5281/ZENODO.3638211>.

⁵ Dawei Lin et al., “The TRUST Principles for Digital Repositories,” *Scientific Data* 7, no. 1 (May 1, 2020): 3, <https://doi.org/10.1038/s41597-020-0486-7>.

⁶ Also available as a Word document under: Alicia Urquidí Díaz and Karen Payne, “HMetS Implementation Plan Template” (WDS International Technology Office, 2021), <https://doi.org/10.5281/zenodo.4589012>.

metadata services. This information package was created by the WDS-ITO to help research data repository managers to create an implementation plan for harvestable metadata services into their data repositories.

Introduction

Since 2016, the FAIR Data Principles—**F**indability, **A**ccess, **I**nteroperability, and **R**euse—defined by Wilkinson et al.⁷ have been guiding good research data management (RDM) practices for the research community. Some organizations such as CoreTrustSeal⁸ and Research Data Alliance⁹ address the community's RDM need to develop requirements and guidelines for data repositories to support FAIR data. Offering machine harvestable metadata to improve the findability of research data is one of the essential requirements for a trustworthy and mature data repository.^{10,11}

Developing machine harvestable metadata services in a data repository can be a complex project, as it depends on a variety of factors, such as the IT infrastructure of the repository, metadata standards, support from the senior management and community, available resources and funding. An actionable plan is recommended to succeed in implementing machine harvestable metadata services in a data repository. HMetS-WG of the WDS-ITO designed an interactive tool that guides research data repository managers step-by-step to create a semi-customized actionable implementation plan for developing harvestable metadata services. The interactive tool is a locally used interactive website with multiple steps and pathways that walk users through the process of creating an implementation plan for harvestable metadata services. In each step of the interactive guide, users will come across some key considerations and relevant online resources to help them make decisions. At the end of the interactive tool, users can print a copy of a semi-customized implementation plan for harvestable metadata services and a list of online resources about harvestable metadata services.

⁷ Mark D. Wilkinson et al., “The FAIR Guiding Principles for scientific data management and stewardship,” *Scientific Data* 3, 160018 (2016): 1–9, <https://doi.org/10.1038/sdata.2016.18>.

⁸ CoreTrustSeal Standards and Certification Board, “CoreTrustSeal+FAIR: Statement of Cooperation & Support,” (October 27, 2020), <https://www.coretrustseal.org/why-certification/coretrustsealfair-statement-of-cooperation-support/>.

⁹ FAIR Data Maturity Model WG, “FAIR Data Maturity Model: Specification and Guidelines” (Research Data Alliance, 2020), <https://doi.org/10.15497/RDA00050>.

¹⁰ Ibid.

¹¹ CoreTrustSeal Standards and Certification Board, “CoreTrustSeal Trustworthy Data Repositories Requirements 2020–2022” (Zenodo, 2019), <https://doi.org/10.5281/ZENODO.3638211>.

The Interactive Tool

The interactive tool is a locally used website, powered by [Twine](#)—an open-source tool for telling interactive and nonlinear stories.¹² Users need to download the [HTML file](#) into a local computer to use the interactive tool in a web browser. All the changes made in the interactive tool are saved in the web browser rather than the [twinery.org](#) website.¹³

There are two ways to use the interactive tool: in a user mode or a developer mode. The user mode allows users to create an implementation plan based on WDS-ITO's defined workflow. No technical skills are required to use this interactive tool in the user mode and we anticipate this will be the preferred method for using this tool by most data managers. The developer mode allows users to customize the implementation plan based on a different workflow according to the repository needs. Technical skills such as HTML, CSS, JavaScript and Macros are required to modify this interactive tool in the developer mode. The tool is under [Creative Commons Attribution 4.0 International license](#), and users are free to share and adapt this work as long as they follow the license terms.

The User Mode

a. Opening the interactive tool

To open the interactive tool in the user mode, open the downloaded HTML file in any web browser. When you open the interactive tool in a web browser, you can see the location of the downloaded HTML file on the address bar of the web browser (Figure 1).

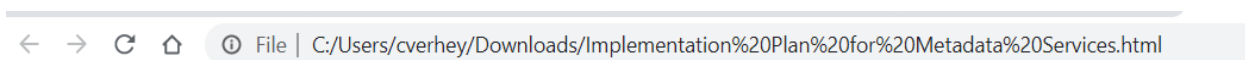


Figure 1: Address bar on a web browser that shows the file location of the interactive tool

The interface of the interactive tool has a sidebar and a passage (Figure 2). The sidebar is located on the left, which contains an overview of this guide, instructions on how to use the site, and the site documentation (i.e., this guidance document). The passage is in the main area

¹² Twine, “Twine” accessed March 16, 2021, <https://twinery.org/>.

¹³ Twine Wiki, “Where Your Stories Are Saved”, last modified October 9, 2017, accessed March 16, 2021, http://twinery.org/wiki/twine2:where_your_stories_are_saved.

of the site, which is the interactive tool that creates an implementation plan for harvestable metadata services.

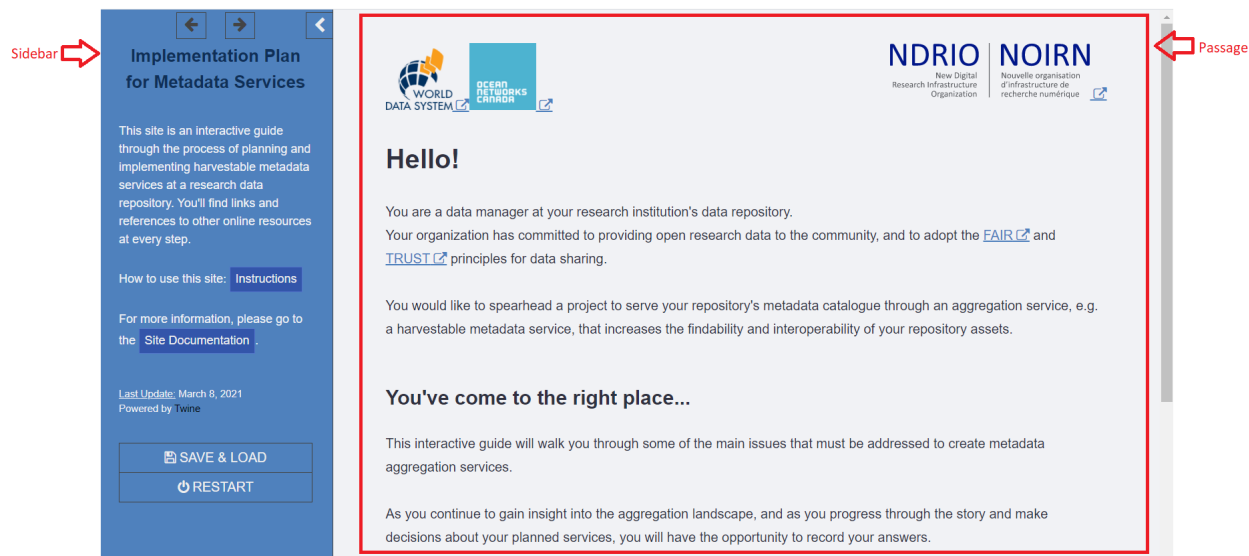
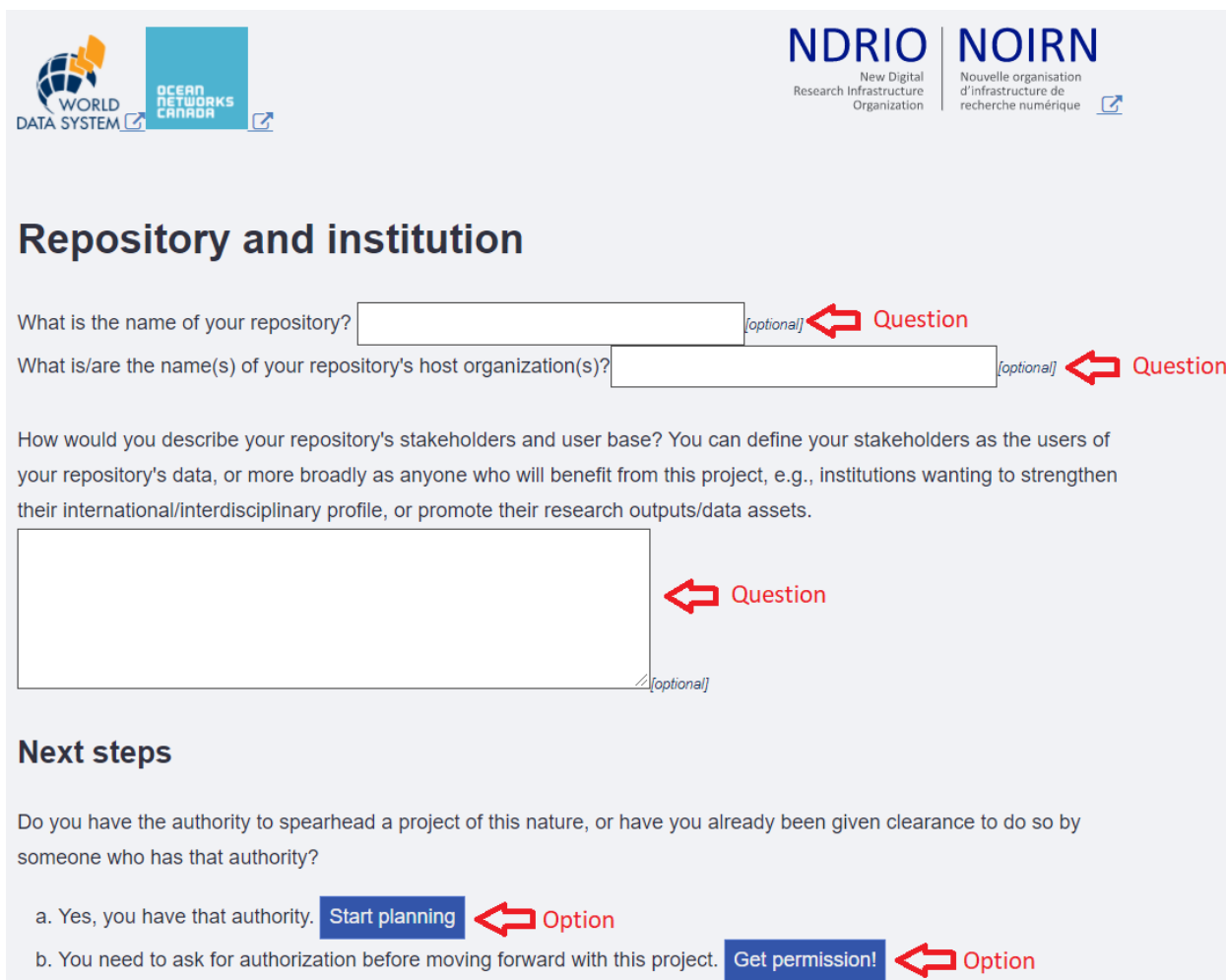



Figure 2: The interface of the interactive tool: a sidebar and a passage


b. Navigating the implementation plan

There is a short story that describes the stage of creating an implementation plan for harvestable metadata services in each passage. Some passages also contain relevant online resources. After reading the short story and/or online resources, a user will come across some questions and/or a few options (Figure 3). It is optional for users to answer the questions in the designated white text areas. The narrative is set up as a sort of “choose your own adventure” series. To move forward to the next step of the implementation plan, a user must click one of the available options. The options are blue buttons usually at the bottom of each passage, although they are at times located at the top of a passage.




Repository and institution

What is the name of your repository? [optional]  Question


What is/are the name(s) of your repository's host organization(s)? [optional]  Question

How would you describe your repository's stakeholders and user base? You can define your stakeholders as the users of your repository's data, or more broadly as anyone who will benefit from this project, e.g., institutions wanting to strengthen their international/interdisciplinary profile, or promote their research outputs/data assets.

[optional]  Question

Next steps

Do you have the authority to spearhead a project of this nature, or have you already been given clearance to do so by someone who has that authority?

a. Yes, you have that authority. [Start planning](#)  Option

b. You need to ask for authorization before moving forward with this project. [Get permission!](#)  Option

Figure 3: The main page of the interactive tool, showing the address bar, sidebar, and passage

To return to the previous passage, a user can click the left arrow located on top of the sidebar (Figure 4).



Return to the previous passage 

Implementation Plan for Metadata Services

This site is an interactive guide through the process of planning and implementing harvestable metadata services at a research data repository. You'll find links and references to other online resources at every step.

How to use this site: [Instructions](#)

Repository and institution

What is the name of your repository? [optional]

What is/are the name(s) of your repository's host organization(s)? [optional]

Figure 4: The return button on the sidebar

To return to the first passage of the implementation plan, a user can click the “RESTART” button on the bottom of the sidebar (Figure 5). However, clicking the “RESTART” button will lose all the unsaved progress. Please refer to the next section, “Saving and restoring your progress,” to learn how to save the progress of the implementation plan.

Implementation Plan for Metadata Services

This site is an interactive guide through the process of planning and implementing harvestable metadata services at a research data repository. You'll find links and references to other online resources at every step.

How to use this site: [Instructions](#)

For more information, please go to the [Site Documentation](#).

Last Update: March 16, 2021
Powered by Twine

[SAVE & LOAD](#)

[RESTART](#)

Repository and institution

What is the name of your repository? [optional]

What is/are the name(s) of your repository's host organization(s)? [optional]

How would you describe your repository's stakeholders and user base? You can define your stakeholders as the users of your repository's data, or more broadly as anyone who will benefit from this project, e.g., institutions wanting to strengthen their international/interdisciplinary profile, or promote their research outputs/data assets.

[optional]

Figure 5: The “RESTART” button on the sidebar

c. Saving/Restoring/Deleting your progress

This interactive tool allows users to save their progress and return to it at a later time. To save a local copy of the progress of the implementation plan (a *.save file), a user can click the “SAVE & LOAD” blue button on the bottom of the sidebar (Figure 6).

Implementation Plan for Metadata Services

This site is an interactive guide through the process of planning and implementing harvestable metadata services at a research data repository. You'll find links and references to other online resources at every step.

How to use this site: [Instructions](#)

For more information, please go to the [Site Documentation](#)

Last Update: March 16, 2021
Powered by Twine

SAVE & LOAD

RESTART

Repository and institution

What is the name of your repository? [optional]

What is/are the name(s) of your repository's host organization(s)? [optional]

How would you describe your repository's stakeholders and user base? You can define your stakeholders as the users of your repository's data, or more broadly as anyone who will benefit from this project, e.g., institutions wanting to strengthen their international/interdisciplinary profile, or promote their research outputs/data assets.

[optional]

Figure 6: The “SAVE & LOAD” button on the sidebar

It will then display a pop-up menu, where a user can select a local directory as a destination to save their progress by clicking the “Save” button (Figure 7).

SAVE & LOAD			×
A	Load	...	Delete
1	Save	...	Delete
2	Save	...	Delete
3	Save	...	Delete
4	Save	...	Delete
5	Save	...	Delete
6	Save	...	Delete
7	Save	...	Delete
8	Save	...	Delete

Save to Disk... **Load from Disk...** **Delete All**

Figure 7: The local directory to save the progress of the implementation plan

To return to the saved implementation progress (a *.save file), users can click the “SAVE & LOAD” blue button on the bottom of the sidebar to open the pop-up menu. Then, users can select any saved implementation progress on the pop-up local directory by clicking the “Load” button (Figure 8).

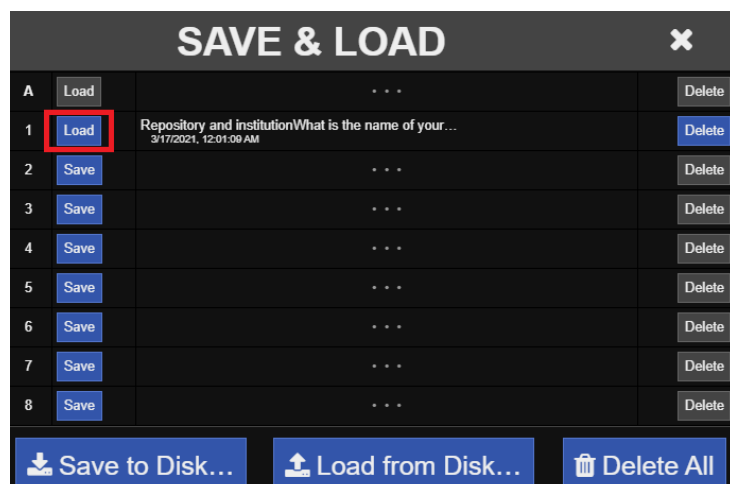


Figure 8: The local directory to load the saved progress of the implementation plan

To delete any saved implementation progress (a *.save file), users can click the “SAVE & LOAD” blue button on the bottom of the sidebar to open the pop-up menu. Users can then delete any saved implementation progress on the pop-up local directory by clicking the “Delete” button (Figure 9).

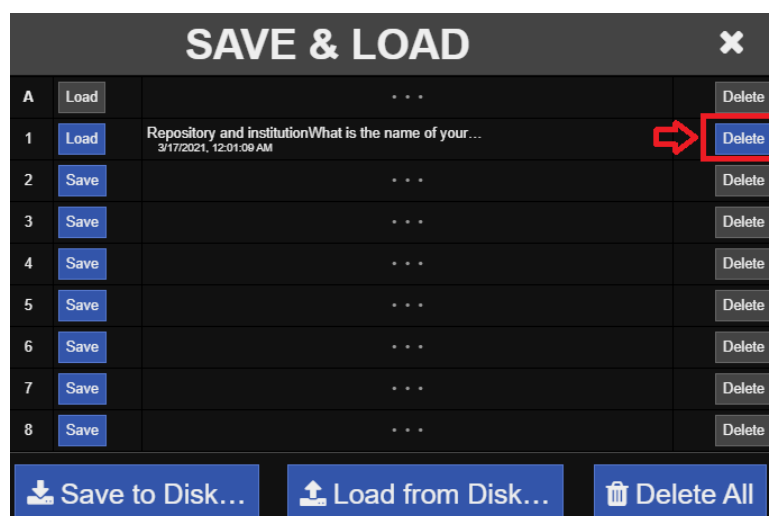


Figure 9: The local directory to delete any the saved progress of the implementation plan

Because the narrative is saved in the browser cache, deleting the browser’s history or changing to a different web browser to open this interactive tool will also delete any saved implementation progress.¹⁴

¹⁴ Twine Wiki, “Where Your Stories Are Saved”, last modified October 9, 2017, accessed March 16, 2021, http://twinery.org/wiki/twine2:where_your_stories_are_saved.

d. Printing the implementation plan

A step before the end of the walk-through, users will come across an option to view and print a copy of the semi-customized implementation plan. All the answered questions and/or chosen options customize the implementation plan for harvestable metadata services. To view and/or print the implementation plan, click the “View and save a copy of your responses (print your customized implementation plan template)” button (Figure 10).

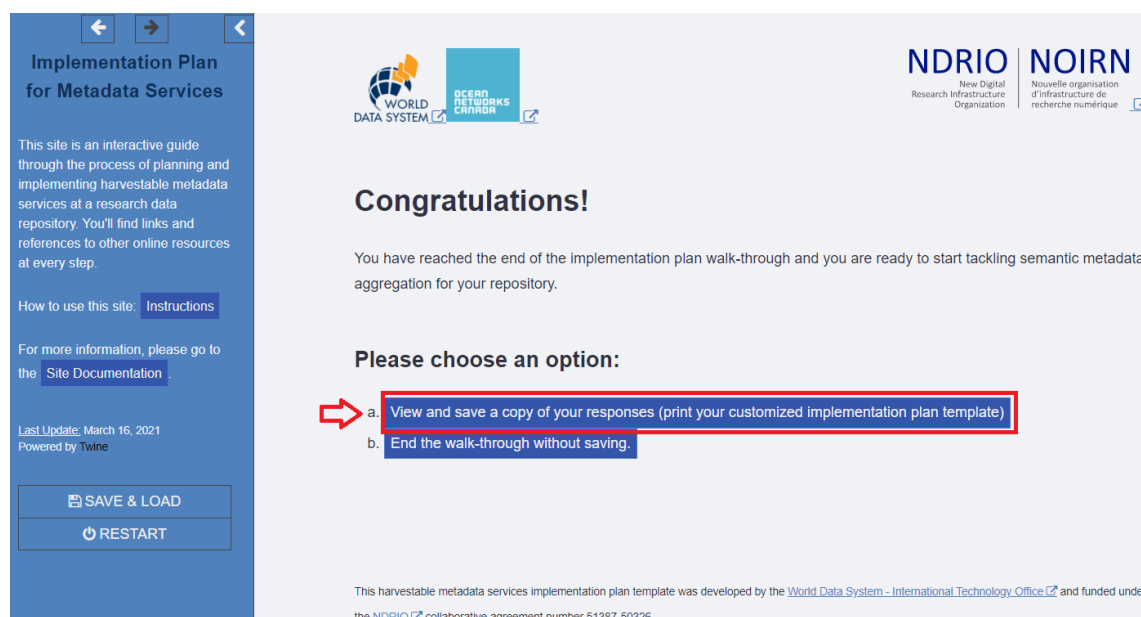


Figure 10: An option to print the implementation plan

Clicking this button will link to a new passage to view the implementation plan. On top of the passage of the implementation plan, there is a “Print your plan” button to print a copy of the semi-customized implementation plan of harvestable metadata services. To optimize the print layout, it is recommended to minimize the sidebar by clicking the < button on the top right corner of the sidebar before clicking the “Print your plan” button (Figure 11).

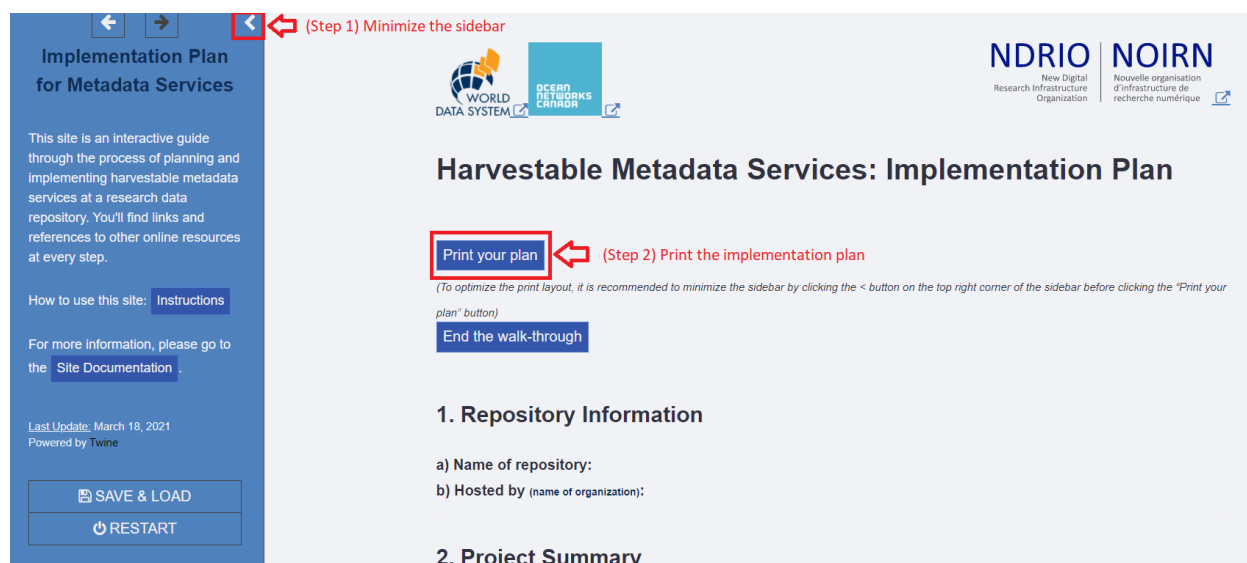


Figure 11: Steps to print the implementation plan

e. Obtaining resources list

After going through the progress of the implementation plan, users will have an option to end the walk-through. Clicking the “End the walk-through” button will lead to the last passage of the interactive tool, which allows users to restart the tool and/or print the resources list. Similar to printing the implementation plan, it is recommended to minimize the sidebar by clicking the < button on the top right corner of the sidebar before clicking the “Print the resources list” button (Figure 12).

The screenshot shows a web page titled "Implementation Plan for Metadata Services". On the left is a blue sidebar with navigation links: "Instructions", "Site Documentation", "SAVE & LOAD", and "RESTART". The main content area has a header with logos for "WORLD DATA SYSTEM" and "OCEAN NETWORKS CANADA", and "NDRIO" and "NOIRN". The text reads: "End of story", "Thank you for playing!", and a "RESTART" button. Below this, a paragraph describes the International Technology Library (InTeLibrary) and provides a link to a Zotero collection. A red box highlights the "Print the resources list" button, with an arrow pointing to it from the text "(Step 2) Print the resources list". Another red box highlights a back arrow in the sidebar, with an arrow pointing to it from the text "(Step 1) Minimize the sidebar". A note at the bottom of the sidebar states: "(To optimize the print layout, it is recommended to minimize the sidebar by clicking the < button on the top right corner of the sidebar before clicking the 'Print your plan' button)".

Implementation Plan for Metadata Services

This site is an interactive guide through the process of planning and implementing harvestable metadata services at a research data repository. You'll find links and references to other online resources at every step.

How to use this site: [Instructions](#)

For more information, please go to the [Site Documentation](#).

Last Update: March 18, 2021
Powered by Twine

[SAVE & LOAD](#)

[RESTART](#)

[Print the resources list](#) (Step 2) Print the resources list

(To optimize the print layout, it is recommended to minimize the sidebar by clicking the < button on the top right corner of the sidebar before clicking the "Print your plan" button)

Resources

End of story

Thank you for playing!

[RESTART](#) to play again.

The International Technology Library (InTeLibrary) collects and shares RDM information resources, most of which are freely available on the web. It is curated and maintained by the International Technology Office (ITO). Visit InTeLibrary's Harvestable Metadata Services collection at https://www.zotero.org/groups/2597955/intelibrary_-_harvestable_metadata_services for more than 300 resources about harvestable metadata services planning, development, and implementation. Areas covered include metadata standards, exchange protocols, metadata aggregation services, certification, assessment frameworks and metrics, as well as data sharing principles, best practices and recommendations.

Figure 12: Steps to print the resources list

The Developer Mode

a. Accessing the developer mode

If you wish to modify the workflow and content of this interactive tool, you need to open the downloaded HTML file in the developer mode. There are two options to access the developer mode: (a) open it in Twine 2 software ([Windows](#), [macOS](#) or [32-bit Linux](#)), or (b) use it online in <https://twinery.org/>. No matter which option you choose, click “Import From File” on the right sidebar of the Twine 2 software or Twinery.org, and then select the downloaded HTML file of this interactive tool (Figure 13).

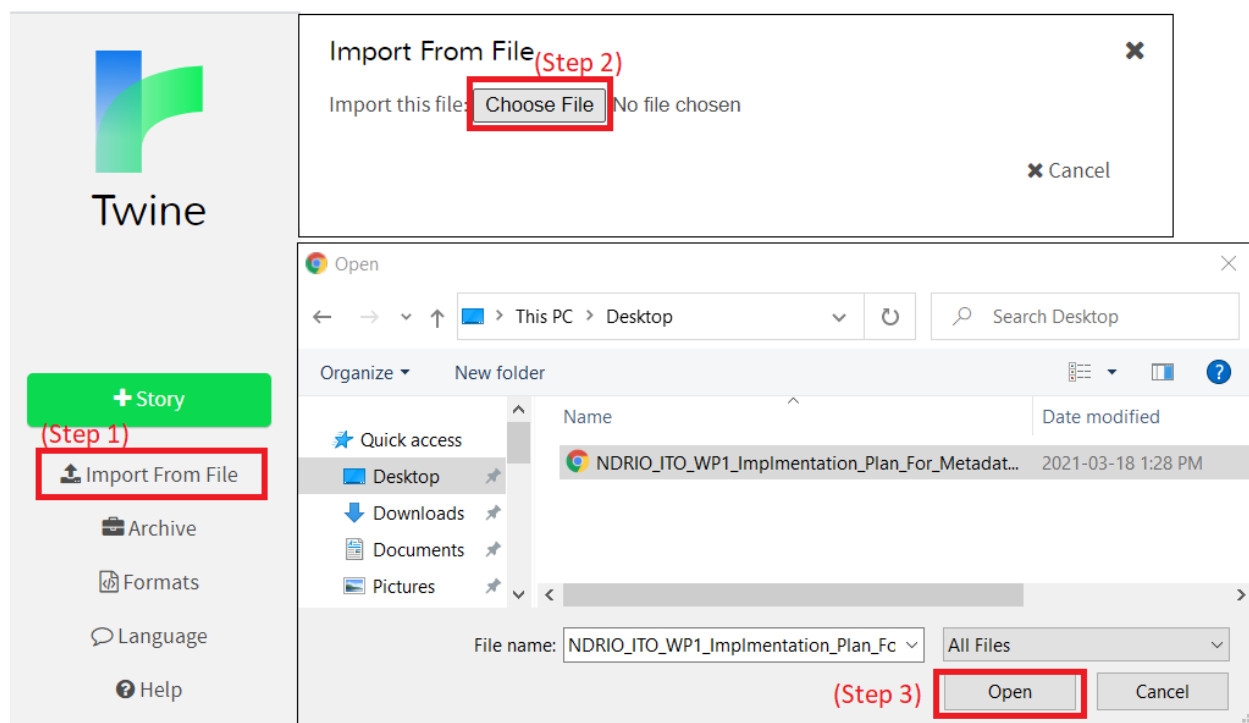


Figure 13: Steps to import Twine HTML file into Twinery.org / Twine 2 software

Once you import the file into Twine 2 software or Twinery.org, you will see the imported story on the main panel (Figure 14).

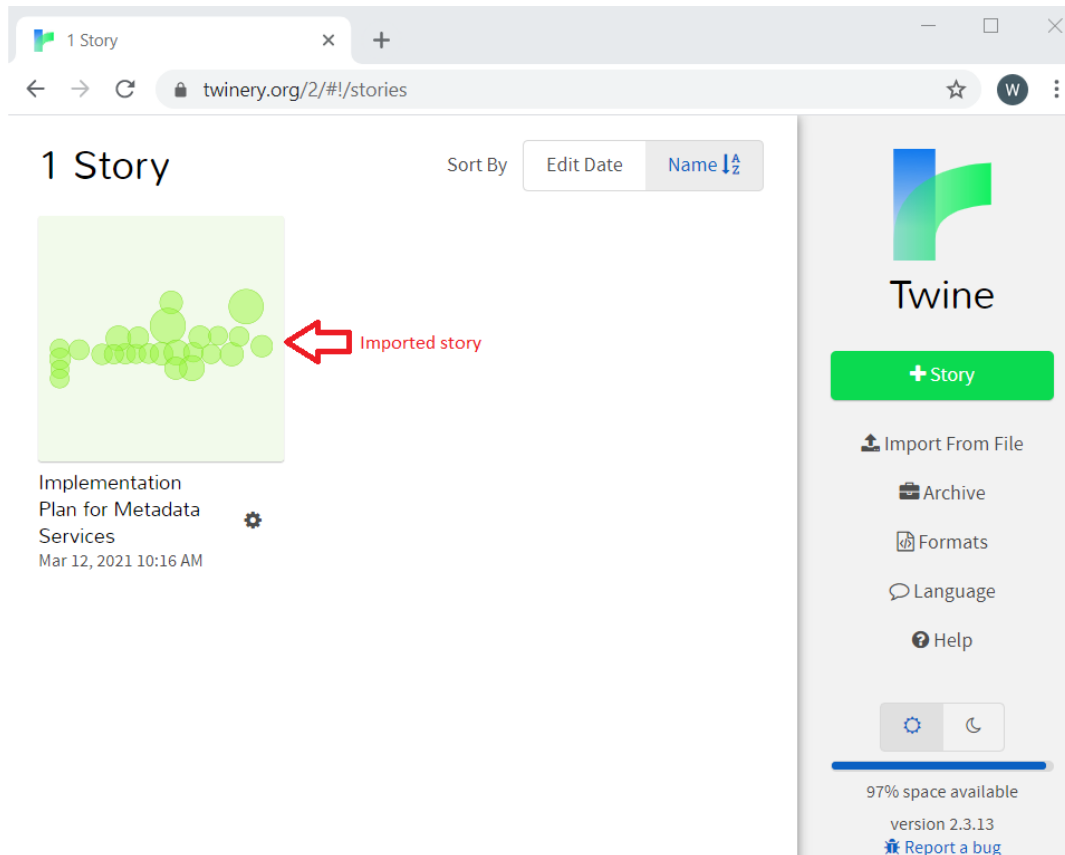


Figure 14: The main panel with the imported file in Twinery.org

Once you click the story, you will open the workflow of this interactive tool. Each block corresponds to a passage or a function. The block named “Start” with a green rocket icon is the first passage of the tool (i.e., homepage of the website). All other blocks with connecting arrows (e.g., “Repository basics,” “Get permission” and “Expected outcomes” passages) are passages of this interactive tool, which contain all the content and interactive functions. The blocks without connecting arrows (i.e., “PassageHeader,” “PassageFooter,” “StoryCaption” and “Instructions” passages) contain content on the header of passage, footer of passage, and the left sidebar (Figure 15).

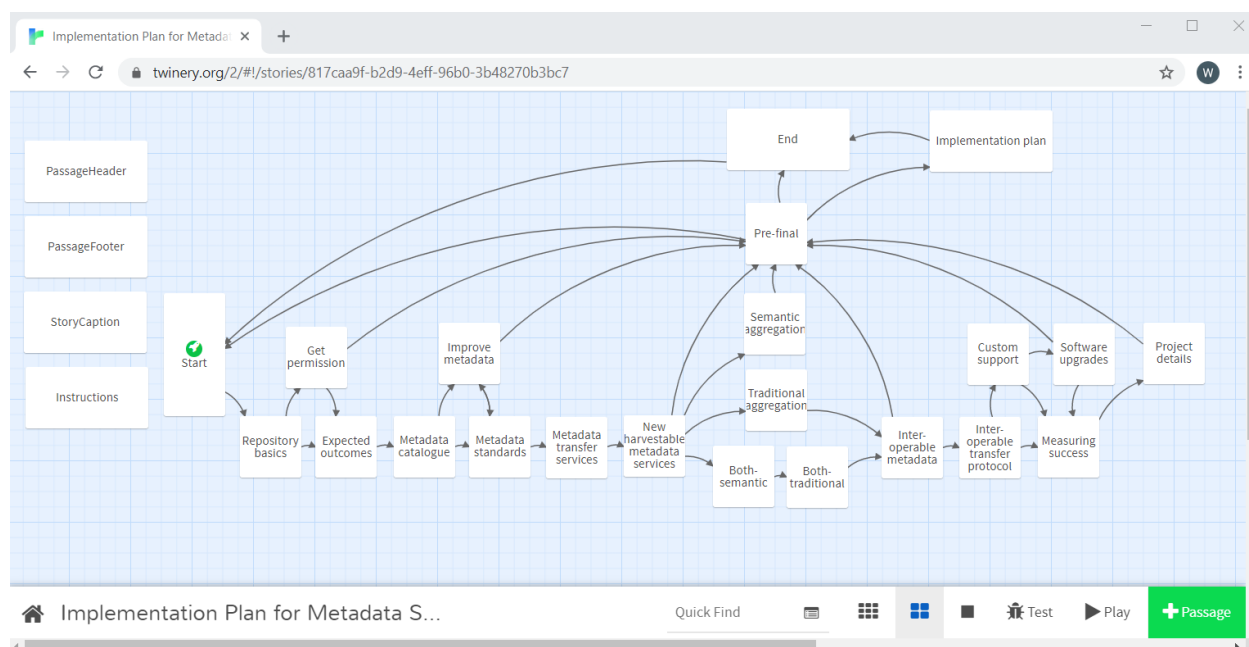


Figure 15: The workflow of this interactive tool

b. Editing the content

Click any of the blocks that you wish to edit. Once you click the block, it will open a menu below the block. Then click the pencil logo (Figure 16) to open an edit panel of the selected passage (Figure 17).

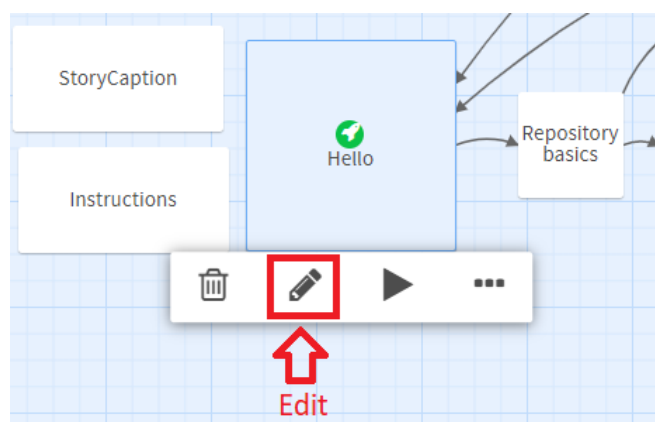


Figure 16: Edit button to edit a passage

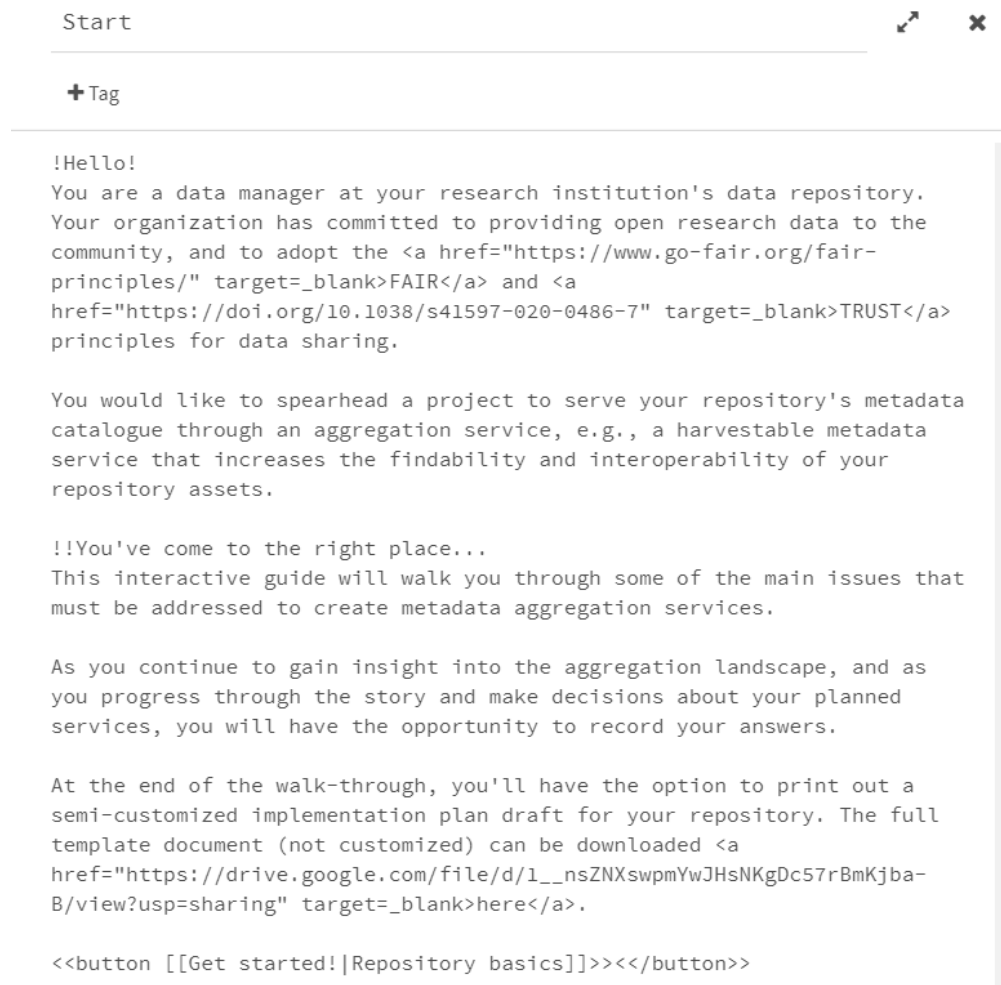


Figure 17: An edit panel of the “Start” passage

You can use plain text, HTML, CSS, Markup, Javascript and/or Macros in the edit panel. Please refer to SugarCube v2 Documentation¹⁵ to learn how to modify a twine file.

c. Saving an edited file

Once you finish editing the twine file, you can save it into a new HTML file onto your local computer. First, click the name of the story on the bottom navigation bar to open a menu. Then click “Publish to File” (Figure 18). It will automatically save a new HTML file of this twine story onto the local computer.

¹⁵ Chapel, “SugarCube v2 Documentation”, last modified March 8, 2021, accessed March 16, 2021, <https://www.motoslave.net/sugarcube/2/docs/>.

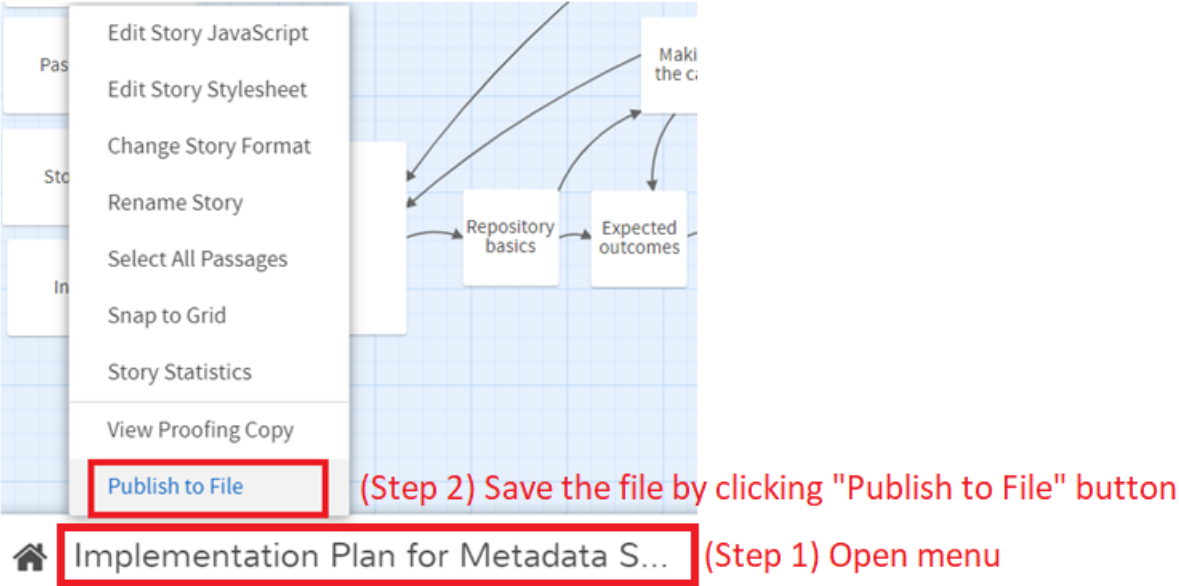


Figure 18: Steps to save an edited Twine file

For details about modifying other functionalities and/or style of the Twine files, please refer to SugarCube v2 Documentation.¹⁶

¹⁶ Ibid.

Limitations and Future Development

This interactive tool is limited to the steps for creating an implementation plan for harvestable metadata services which are pre-defined by HMetS-WG of the WDS-ITO. Although HMetS-WG has incorporated a variety of relevant resources and use cases from WDS members to create this interactive tool, this tool may not fit the situations of all the data repositories. Users may require using the developer mode to customize the implementation plan that fits their needs, which may be a barrier to some research data managers who may lack resources and/or skills to use the developer mode. This tool may need to be modified from time to time to meet the community's RDM needs.

Due to constraints at WDS-ITO, we are releasing the first version of the interactive tool as an HTML file that requires users to use it on their local computer. Future development of the interactive tool will depend on the uptake, demand and use of the data by the RDM community. The WDS-ITO has plans for a website revamp in 2021 and will assess the appropriateness of including the tool as part of the website.