

Developing an Interactive Data Deposit Decision Tree

University of Toronto: Leslie Barnes, Digital Scholarship Librarian; Dylanne Dearborn, Research Data Management Coordinator; Jasmine Lefresne, Data Librarian; Steve Marks, Digital Preservation Librarian

OBJECTIVE

Most existing resources that guide researchers to appropriate repositories in which to deposit their research data filter selection through a specific lens, such as funder requirements, FAIR principles, or sensitive data. Our Decision Tree will be an interactive tool that allows users to choose the best path to repository selection based on their own priorities. Three objectives guided us:

- Our Decision Tree needed to account for institutional and national contexts, including funder requirements, legal and ethical frameworks, and institutional practices;
- We wanted to help researchers make informed decisions about depositing their data and cultivate a deeper understanding of the many factors that might impact their choice;
- We wanted to develop a flexible and interactive tool that responds to researchers' varying priorities and motivations, while respecting their time.

The Decision Tree allows researchers to choose their own starting point. Each starting point leads to different question ordering, ensuring users see the most relevant questions first.

HOW IT WORKS

Our Decision Tree combines two models in one easy-to-use interface called the Repository Selector. The first model mimics the conversational flow of a consultation. We developed user personas that represent researchers' different motivations and created sequences of yes/no questions. The second model helps users navigate generalist repository options based on criteria that operate as facets. This approach allows us to ensure users are choosing the best repository for their data.

BENEFITS

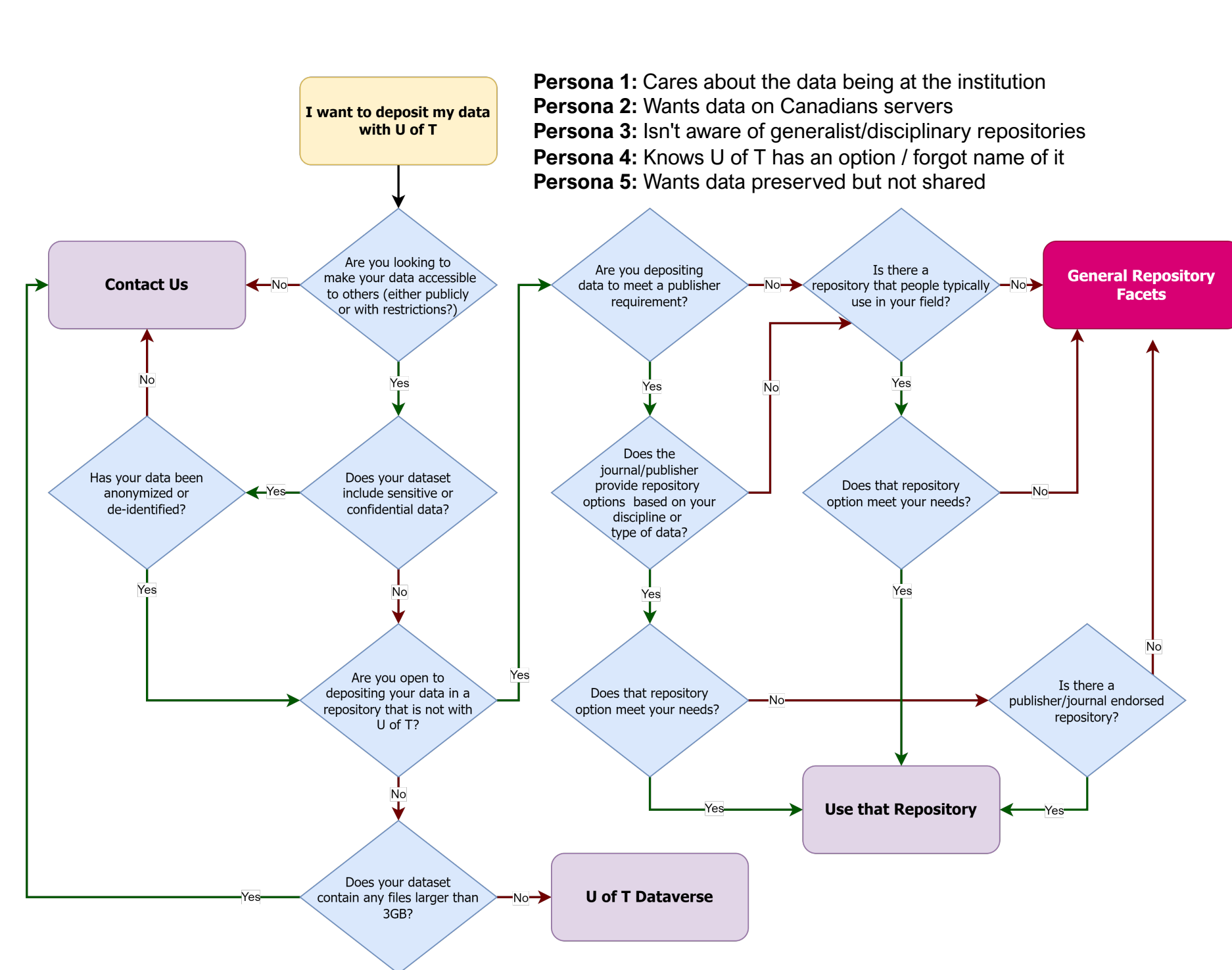
Promotes Open Science: A selection tool encourages data sharing and depositing, reduces decision-making barriers, and introduces ideas of reuse and discovery while respecting disciplinary norms.

Flexible and Interactive: We will adapt the Repository Selector as needs evolve, introducing new starting and end points, and modifying question flows as needed.

Expands Reach and Efficiency: As an online resource, the Repository Selector will always be available to all users, reducing consultation burdens and allowing users to drive their own decisions. It is also a valuable tool for those who prefer self-service.

Standardizes Consultation Approach: The Repository Selector will facilitate training for new librarians and standardize conversations across multiple campuses. It can also generate context for any subsequent in-person consultations, as the researcher will have already engaged with some of the concepts and their initial output path can be attached to the consultation request.

Provides Institutional Insights: Analysis of user paths, motivations, and choices will generate insight into institutional practices, highlight gaps in services or infrastructure, and provide valuable data for future development.



Persona 1: Cares about the data being at the institution
Persona 2: Wants data on Canadian servers
Persona 3: Isn't aware of generalist/disciplinary repositories
Persona 4: Knows U of T has an option / forgot name of it
Persona 5: Wants data preserved but not shared

STARTING POINTS

- I want to share my data so people can find/use it
- I need a DOI for my dataset
- I want to deposit my data with U of T
- I need to deposit my data to meet my funder requirements
- I need to deposit my data to meet a journal requirement
- I need to include a repository in my DMP
- I want to deposit my data but only share it with certain people
- I want to deposit a very large dataset
- I want to deposit confidential or sensitive data

CONSIDERATIONS AND LIMITATIONS

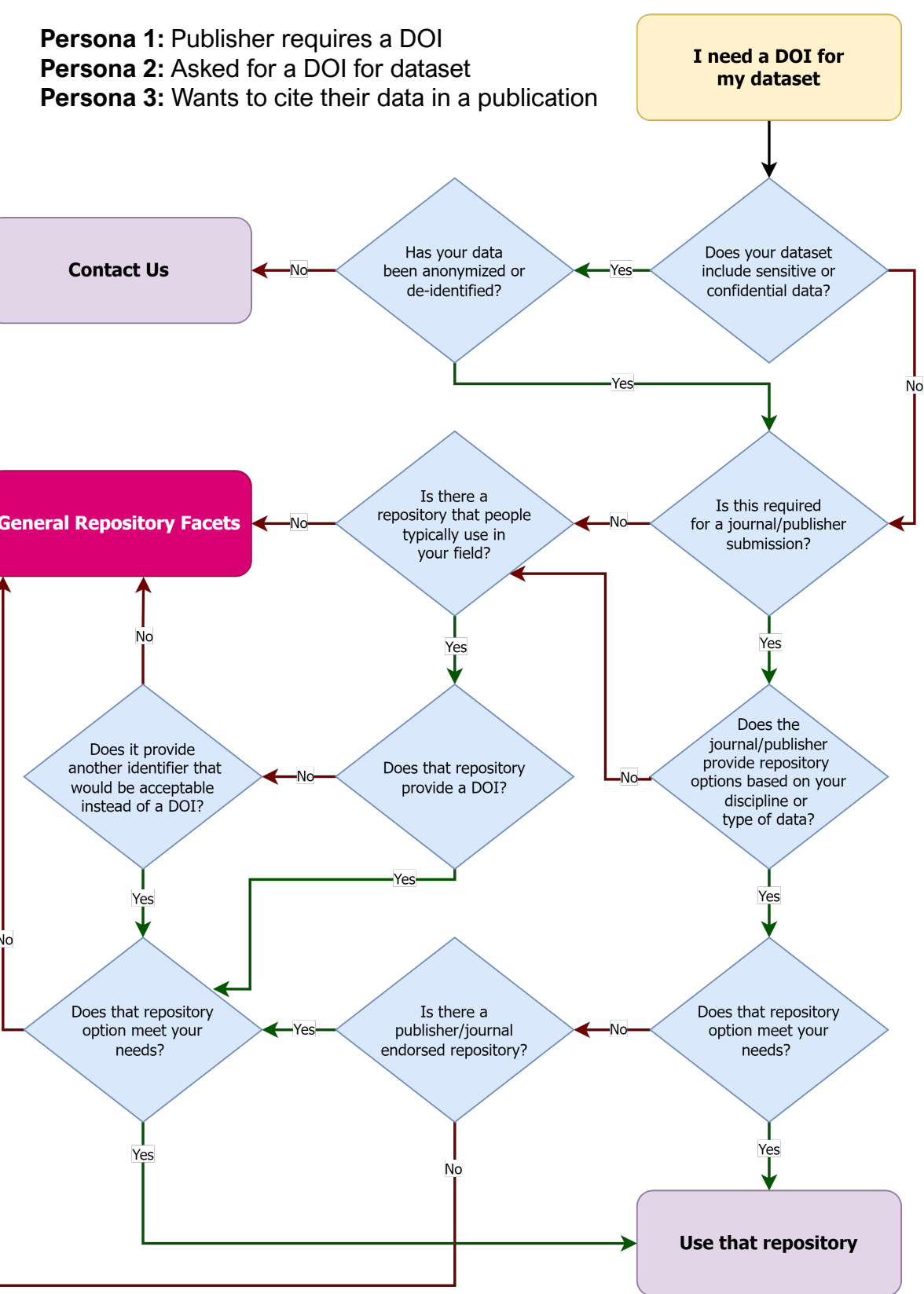
Subject Repository Inclusion: The vast number of subject-specific repositories prevented us from listing individual ones in the Repository Selector. One option is to link to recommended lists, though this limits the granularity of choice offered to researchers.

User Interpretation: Path effectiveness relies on accurate and complete information. Since users may have minimal experience with data deposit and repository selection, we've created built-in guidance to provide context and minimize confusion.

Oversimplification: We cannot fully replicate a consultation and therefore risk not capturing all nuances in decision-making. Additionally, the tool can't represent all user needs or scenarios.

National requirements TBD: The Tri-Agency has yet to release a RDM Policy for Data Deposit. Once announced, we anticipate adding starting points and altering question flow and recommendations.

Maintenance and Updating: To stay current, the tool will require continuous updates based on user feedback, and evolving requirements and best practices. We will also need to monitor included repositories and update the tool as needed.



Persona 1: Publisher requires a DOI
Persona 2: Asked for a DOI for dataset
Persona 3: Wants to cite their data in a publication

POSSIBLE END POINTS

- | | | | |
|-------------------------|-------|----------|------------|
| Disciplinary Repository | FRDR | figshare | Zenodo |
| U of T Dataverse | Dryad | OSF | Contact Us |

NEXT STEPS

The next steps for the Repository Selector development include developing more starting points for the question flow and adding more embedded guidance. We also need to finalize the supporting resources, including the "Repository Evaluation Criteria" and the "Maximizing Data Impact" guides.

We are also assessing the feasibility of creating and maintaining curated lists of disciplinary repositories. A key priority is to develop a way to capture and export decision paths to facilitate subsequent consultations. In addition to bridging between self-service and in-person consultation, this ability will help us better understand behaviour and inform future developments. After finalizing questions and content design, we will move forward with the technical implementation and usability testing with our research community.

WORKS CONSULTED

ARDC. (2021). Publishing and sharing sensitive data - flowchart. Zenodo. <https://doi.org/10.5281/zenodo.7259737>

Canadian Association of Research Libraries & Portage. (2019). Repository Options in Canada: A Portage Guide. https://library.ulethbridge.ca/ld.php?content_id=34822303

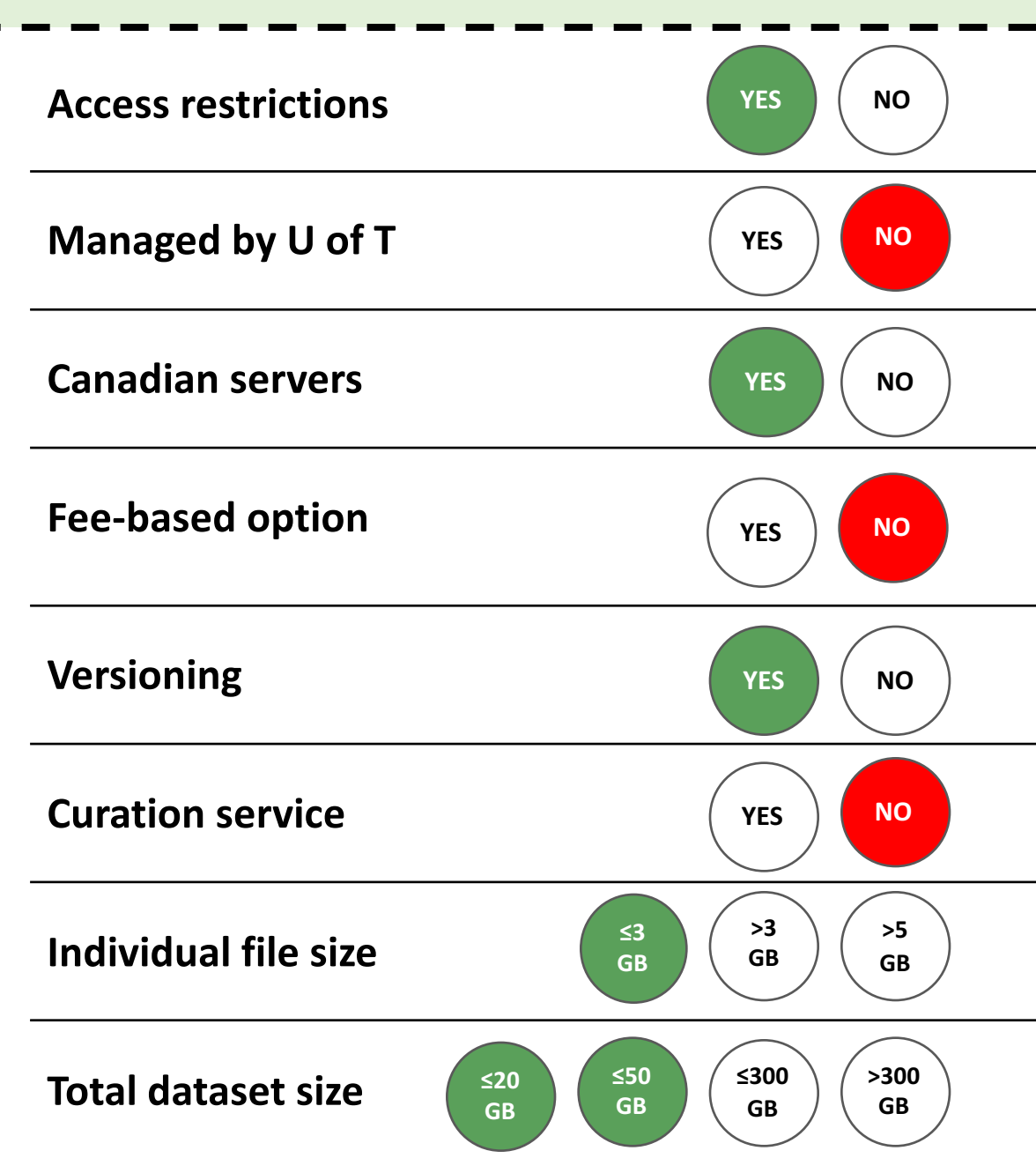
Enabling FAIR Data Community, Ruth Duerr, Danie Kinkade, Michael Witt, & Lynn Yarmey. (2018). Data Repository Selection Decision Tree for Researchers in the Earth, Space, and Environmental Sciences. Zenodo. <https://doi.org/10.5281/zenodo.1475430>

Utrecht University. (2023). Data Repository Finder. <https://tools.uu.nl/repository-decision-tool/>

INTERFACE MOCKUP SOURCE CODE

Creative Commons. License Chooser beta, Creative Commons. <https://chooser-beta.creativecommons.org/>

GENERAL REPOSITORY FACETS



INTERFACE MOCKUP

