



We Can Work It Out

Cross-Functional
Collaboration on Repository
Strategy at New York
University

Open Repositories 2024 June 6



Working group charge

Identify a number of options related to merging or maintaining separate services based on a set of needed features and functionality. For each option, assess the labor involved for related activities, such as migration and implementation, as well as the risks associated. Consider repository services from both the maintainers and user experience perspectives. **Make a recommendation for whether or not to merge the IR and data repository functions into a single service.**

Smaller teams for research activities

Metadata

Internal Research

Current Services

Functional Requirements

Peer Benchmarking

Software Resource Assessment

Software Governance

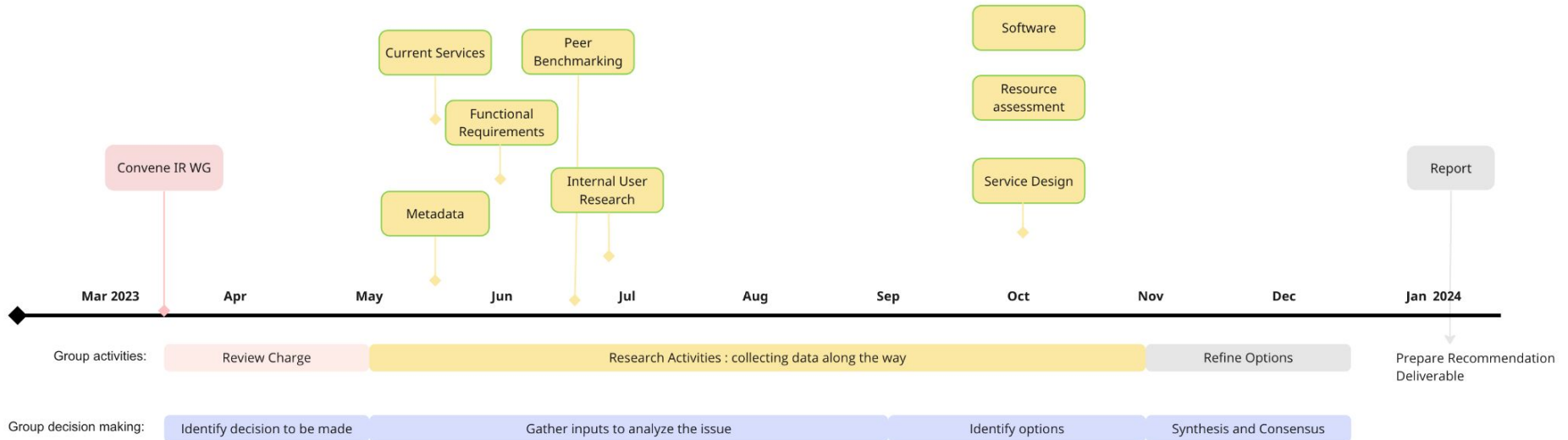


If you are willing to lead a task or related group of tasks, **bold** your name.

What is the ef

#	ACTIVITY		WHO	SIZE
1.1	Perform a lightweight metadata assessment of what is in FDA <ul style="list-style-type: none"> Review metadata registry for unique terms Review input forms config file for custom collection schema 	Metadata	DV, TC	Small
1.2	Document metadata schema for UV <ul style="list-style-type: none"> Data Cite schema Note any local custom fields 	Metadata	DV	Small
2.2	Create functionality matrix: current, desired, overlaps <ul style="list-style-type: none"> Document the functionality for UV, FDA (and SDR??) Maybe base it on the original DRSR matrix for consistency Maybe reference the newer Federal requirements for 'desirable characteristics' 	Functional requirements	DV, CK, ZC, JG	Large - Requires more organization and time to complete
2.3	Align repository goals & capacities with DoL digital preservation vision <ul style="list-style-type: none"> Read the DoL policy statement posted on DoL website Does FDA and UV meet this vision? 	Functional requirements	DV, TC	
3.1	Look at what our peers are doing in this space (IR, DR, SDR)	Peer benchmarking	CK, AC	Medium - depends on # of peers
3.2	Map COAR report to our offerings - find new functionality	Peer benchmarking	TC, VR, LH	
3.3	Identify and talk with DCN peers who run multiple repos.	Peer benchmarking	TC	
4.1	Assess our technological capabilities and resource availability	Resource assessment	CK	Small
4.2	Assess long term resource requirements for potential models <ul style="list-style-type: none"> what might staffing for service team and development team look like for two repositories running on the same software? For one combined repository? For two repositories running on different software (the status quo)? How does being active users/participants of two open source projects affect our time commitment? 	Resource assessment	CK, NW	Medium
4.3	Identify staffing and skill capacities required for desired service model(s)	Resource assessment	CK, NW	
5.1	Access and read UV service documentation	Current services	AC, ZC	

Collecting data along the way



Consensus on what is most important

The most important things that we need to keep in mind when we make our recommendation are

Keep the user at the center

Patron convenience

Ease of use for end user

Keep many user groups in mind

Taking into account user's future needs

Sustainability : design something that will remain well used

Impacts on curation and workflows

Clarity on curation responsibilities

Smooth workflow for staff with clearly defined roles and responsibilities

Feasibility of decision

How hard it is to migrate content

Realistic to timeframe

Ease of long term support

Sustainability of platform/service

Technical maintenance

Support humans working on service

Emphasizing people as service (not just tech)

That developers are excited to work on it

Team collaboration and info sharing

Impact of staff size on ambitions

Supporting consistent branding

Branding and service clarity for users

Supporting consistent branding

Missions aligned

Alignment with mission for repository services

Supporting Library's priorities

Meets functional requirements

Process for digital preservation

Different file types

Persistent links

Draft records for sharing

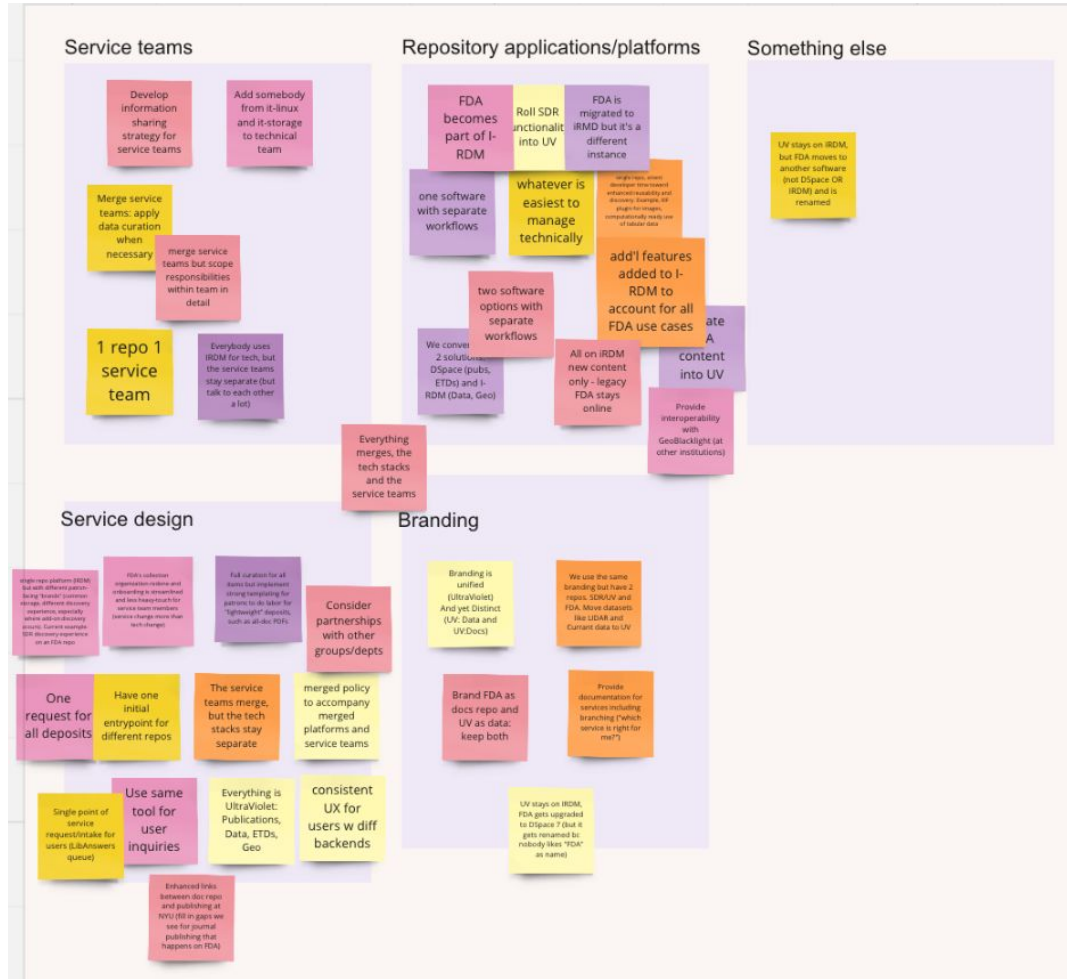
Different use cases for publishing

Meeting must-have functional requirements

Visioning: What are the options?

Individual responses were grouped with “similar”.

Four categories emerged.



What is our current reality?

Resource assessment for the software related options

High level comparison of functionality in DSpace (IR), InvenioRDM (Data Repository) and homegrown Spatial Data Repository



APPENDIX C: IRWG Subgroups

Software Options Advantages and Disadvantages

OPTIONS TOGETHER

One of these

1a FDA becomes part of I-RDM (one instance)

1b FDA is migrated to I-RDM but it's a different instance

AND

One of these

2a SDR fully merged into I-RDM and GeoBlackLight discontinued

2b SDR bitstreams migrated to I-RDM and keep GeoBlackLight as Geodata Discovery Platform

Advantages:

- We can focus our energy on developing one platform.
- We can make a big impact on the open-source community.
- Users are directed to one service platform for repository publishing.
- I-RDM is designed for the complex, multiple object types of data we expect in the future.
- Peers are interested in merging IR with their I-RDM data repositories, notably CERN.
- Allows for a more streamlined service design with a common set of procedures in either separate services or a combined service.

Resulting report

The working group recommendations:

Merge the IR and data repository functions into a single service running on InvenioRDM.

Merge the service teams.

The group reviewed a draft of the final report deliverable in real time, seeking consensus before closing.

01. Repository Applications Recommendations

02. Service Design Recommendations

03. Service Team Recommendations

04. Branding Recommendations

Thank you

Deb Verhoff
dv47@nyu.edu

Digital Collections Manager
New York University Libraries

