DRS (Digital Repository Services) Futures Open Meeting Summary

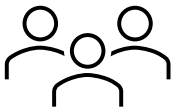


Table of Contents

Executive Summary

The DRS Futures Team is in the Discovery phase of determining the next-generation digital preservation repository. As one of many opportunities to engage everyone within the Harvard community, the DRS Futures team invited all departments within Harvard to participate in the Open Meeting. Participant insights inform the ideal repository and shape the requirements and outcomes. Participants expressed their preferences, goals, and requirements for the new repository. Detailed insights about workflow, functionality, and needs were shared by participants and informed the observations of this summary, however, common themes emerged in all aspects of the open meeting. Participants consistently voiced that the new repository needs to be:

- Easy to use
- Seamlessly integrated with other Harvard systems
- Able to support all digital formats
- Scalable and flexible

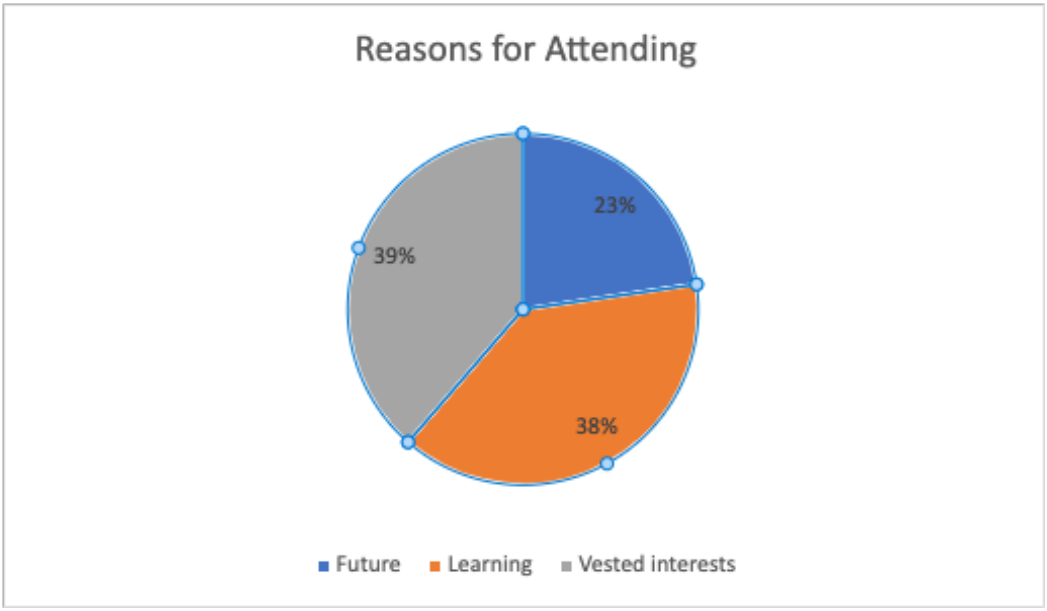
More insights and trends are captured with greater nuance in the sections below. The DRS Futures team will offer more opportunities to discuss needs, preferences, and goals throughout the Discovery Phase.

Open Meeting Details

The Digital Repository Services (DRS) Futures Open Meeting was held on January 25th at 3:00pm Eastern time. The entire Harvard community was invited to attend, including the Harvard Library, all the professional school libraries, museums, archives, and anyone in the greater Harvard community. There was strong interest – 237 individuals registered for the event and 126 signed on to take part.

Why Did Participants Attend?

The meeting opened with participants sharing what motivated them to attend. Participants were inspired to attend the open meeting to learn more, had an interest in the future of Harvard's preservation repository, or had a personal stake in how the repository functions and wanted to make sure that viewpoint was represented.



Graph 1. Reasons for attending the open meeting.

The DRS Futures team presented the project plan during the 60-minute, remote presentation.

Breakout Brainstorming Sessions

The participants were grouped into small breakout discussions and took part in a Balloon/Ballast brainstorming exercise in which they were asked to name characteristics that would have a positive impact on the future repository (the balloons) and characteristics that would impede the future repository functions (the ballast). Participation in these breakout sessions allowed the DRS Futures team to clearly discuss with the entire Harvard community the elements of an ideal repository which will be used to shape the requirements and outcomes for DRS Futures. Participants were invited to create lists of beneficial characteristics and challenges to address in the next Harvard repository iteration.

A common thread amongst all the comments and discussions was that participants all felt the repository needed to be easy to use. From a clean, customizable user interface to simplifying and supporting all aspects of data preservation, ease of use was a constant comment from participants in every breakout group. Participants wanted a drag-and-drop deposit experience; an intuitive system that they don't notice much as users; easy metadata editing; a straightforward way to get content out of the repository; and improved searchability. In every category discussed, participants imagined a system that was easy to use and easy to integrate into their various workflow requirements.

Observations and goals for the new repository fall into the following categories:

General Repository Characteristics

- **Nimble and flexible**: participants want a repository that can adapt to future needs and formats

- **User-friendly**: participants want a repository that makes their work easier and more efficient with a clean, intuitive interface.
- **Secure**: participants value the security of Harvard's digital content and want a repository that can address the long-term security of the data.
- **Maintainable:** it is important to the Harvard community that the new repository be able to be supported and updated.
- **Well-documented:** participants would like the new repository to have clear, up-to-date documentation and a plan for keeping the documentation current.

Longevity

The DRS gets top marks for preservation of data. Several participants mentioned this as a strong positive that needs to continue in the next generation of the repository. Participants commented on the repository's value, supplying dependable longevity for digital content for departments within Harvard and potential donors and creators of content. It was also mentioned that Harvard University support increases their confidence in the repository – they were specifically pleased that the DRS was supported institutionally and felt that gave the users and donors greater confidence in the continuation of the repository and the preservation of the data than had it been supported by a single department.

Interoperability and Seamless Integration

This category of comments highlighted participants' strong desire for the new repository to integrate its functions with other Harvard systems to support departmental workflow requirements. One participant commented that they hoped the repository "will be so integrated into workflows, systems, ways of working, that it won't seem like a repository except for those who need it to." Interoperability came up almost as often as ease of use in the breakout discussions and there were frequent comments about how the future repository should integrate with other systems.

User-Friendly Deposit

Participants in the breakout sessions found that a user-friendly deposit was a strong concern. They would like to see a drag and drop deposit choice. They would also prefer that data formatting issues be highlighted early in the deposit process rather than having it come up later in an error email.

Robust and Extensible Metadata Creation

There was a robust discussion about metadata with preferences for robust, extensible metadata creation. Several participants were interested in batch editing metadata or editing metadata at scale. Others were interested in metadata creation supported by artificial intelligence (AI) and machine learning (ML). Participants were interested in file structures and the contextual relationship of files being supported by metadata. There was also interest in being able to pull metadata from other systems and input that into the repository.

Access Management

Participants had specific interests in an account management system that would support the ability to change access flags on digital content on demand. There is concern that some files or objects need to have fine-grained permissions possibilities beyond the ones currently allowed to support researchers both within the Harvard community and more broadly. Participants imagined a system in which they could automate or schedule future actions as copyright status changed, for example, "Notify staff in 10 years that the access flags for objects X-Z should be reviewed." Automated bulk changes to access flags for entire collections rather than files and objects was also mentioned.

Low Cost; Low Barrier to Entry

Participants commented that their *ideal repository should have a low barrier to entry, low storage costs, perhaps supported by an endowment or other dedicated funding.* Participants were concerned that it isn't always clear what digital content is considered institutionally valuable. Several participants discussed the costs of using a repository system as something they would like addressed in the future. Participants mentioned that the cost of DRS is a barrier to entry for departments as content without billing codes can't go into the repository. Participants were concerned that it isn't always clear what digital content is considered institutionally valuable.

All Formats Supported

Multiple participants mentioned they would like to see *the future repository accept and support all types of files and have the flexibility to accommodate new content types and content models.* A/V files and architectural files were mentioned as examples; however, the consensus was that the new repository should be format agnostic.

Infinitely Scalable

Participants want the new repository to be adaptable to an increase in scale of both file sizes and number of files with the goal of "infinite scalability."

Analytics on preservation data

Analytics were an important component of the discussion, however, participants conflated preservation analytics with end-user analytics in some discussions. The future repository needs to integrate with systems that provide end-user analytics. The future repository needs to support detailed, transparent preservation analytics. Participants expressed interest in transparency in the repository for depositors and content managers about fixity – "I trust the DRS, but I'd like to be able to see the checksums." It will be beneficial to have clarity on the kind of analytics the new system will be able to provide.

Accessibility

Participants wanted an accessibility-focused repository with structured formats.

Policy and Eligibility

Participants feel like they need help in answering the question: "where do we put all the stuff we are generating?" They would like a policy clearinghouse that analyzes the most effective/efficient storage options for content or a transparent Harvard-wide policy about what content is eligible to go into the DRS.

The Positives Outweigh the Negatives

Participants were also invited to comment on challenges that keep the repository being ideal. Largely these comments were the mirror images of the wishes and goals discussed above. Comments focused on the difficulty of using systems that weren't as well integrated as they could be, the lack of APIs, inability to batch edit metadata, batch ingest content, or batch download, and the dissatisfaction with limited and inflexible content models. Overwhelmingly the breakout sessions created extensive lists of positive, desirable characteristics, showing that the Harvard community is enthusiastic about the possibilities that can be supported by the new repository. They are imagining big things and excited for the future repository.

Beyond the Repository

Some recommendations supply important insight as to what users are looking for even though they do not directly apply to the repository itself. These comments are noted and reported for the insights they provide about participant needs. The out-of-scope comments are in three categories:

Access and Delivery

This is the largest category of out-of-scope comments. This highlights the need for users to have a well-integrated system that provides them with access to materials they need – as well as the need for the future repository to be flexible enough to integrate with a wide variety of access options.

Search and Discovery

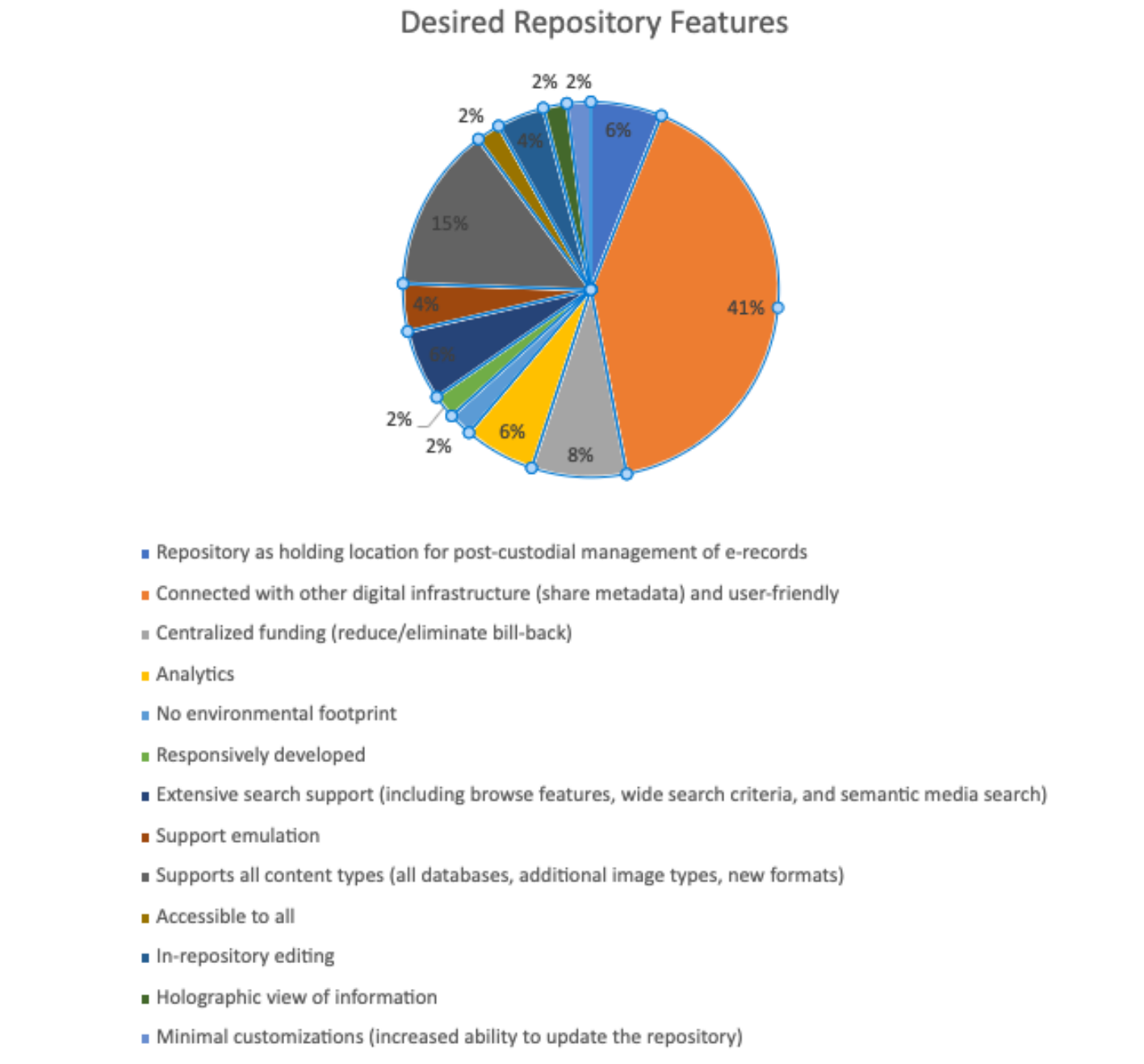
Curators using the administrative user interface for the repository will be supported. However, participants also commented on their own end-user needs for search and discovery of all digital content which is still a role of the delivery services that supply access copies (e.g., HOLLIS). There was some concern that currently end-users placing their data requests are asking the wrong departments for access to digital content.

Analytics on Access/Delivery end-user data

Participants conflated the preservation analytics and those on how end-users have engaged with digital content. Outside of the scope of the repository analytics, users would like to have meaningful usage statistics, impact statistics, cited resources, commercial impact, joy impact, published ORICD records, analytics that feed into Faculty360, generate CVs, and support tenure review.

Parting Wishes from Participants

The meeting concluded with an opportunity for participants to offer a parting wish about the features of the future repository.



Graph 2. Parting wishes for desired repository features.