



Community Meeting

<https://www.dataone.org/meetings/2020/>

Zoom: 943 0955 6242 Pwd: 3282663

One tap dial in: +16699006833,,94309556242#

[code-of-conduct](#), [@DataONEorg](#)

---

## Session 2

### Emerging services from the DataONE Network

⤴ Return to [README](#) : ⤴ Back to [Session 1](#) : Forward to [Session 3](#) ⤵

#### Resources:

Menti: <https://www.menti.com> Code: 62 15 81

DataONE Services page: [dataone.org/services](https://dataone.org/services)

Breakout session notes:

[Group A](#) - DataONE Plus

[Group B](#) - Hosted Repositories

[Group C](#) - Consulting Services

[Group D](#) - Learning Services

#### Notes:

- New services from DataOne
- Core - Provenance, Semantics, Replication, Metadata quality continued

New DataOne Plus features:

DataONE Plus Portals

- DataONE Plus Portals - search across custom data collections, build custom content, create custom search filters
- Portals still pull from the same DataONE repositories, but allow for more customization
- Brand a portal to match your own infrastructure, web presence
- Good way for a project to organize data and provide web presence
- Can provide collection specific overview of how much someone's data collection is being accessed and views, plus FAIR scores for the portal data collection
- Portals are long-term archived, the same way as scientific data
  - Compare to another website option, like Wordpress, which requires maintenance forever, even after project funding ends
- Custom search filters - will be able to add them in the portal builder in a future release, though it requires quality metadata

Hosted Repositories



## Community Meeting

<https://www.dataone.org/meetings/2020/>

Zoom: 943 0955 6242 Pwd: 3282663

One tap dial in: +16699006833,,94309556242#

[code-of-conduct](#), [@DataONEorg](#)

---

- Hardware (1TB of storage), software (mature & stable, web client), and architecture provided by DataONE, data curation carried out by organization
- Hosted Repositories include unique DOIs, access to DataONE Plus portals, data replication & preservation
  - Allows your data to be available even when repository might be offline for maintenance or outage
- Examples: [California Ocean Protection Council \(OPC\)](#) and [Dangermond Preserve](#)

### Consulting Services

- Example uses: Extending web portals, automating metadata collection, metadata review & analysis by trained teams
- Open source platform - anyone is welcome to contribute!
- Permafrost discovery gateway - example of some of the customization that's possible

### Learning

- Skillbuilding Hub
- Short courses

### Services & Pricing

- Portals are in a preview period
- <https://www.dataone.org/services/>
- DataOne Plus - Subscription depends on size of organization
-



Community Meeting

<https://www.dataone.org/meetings/2020/>

Zoom: 943 0955 6242 Pwd: 3282663

One tap dial in: +16699006833,,94309556242#

[code-of-conduct](#), @DataONEorg

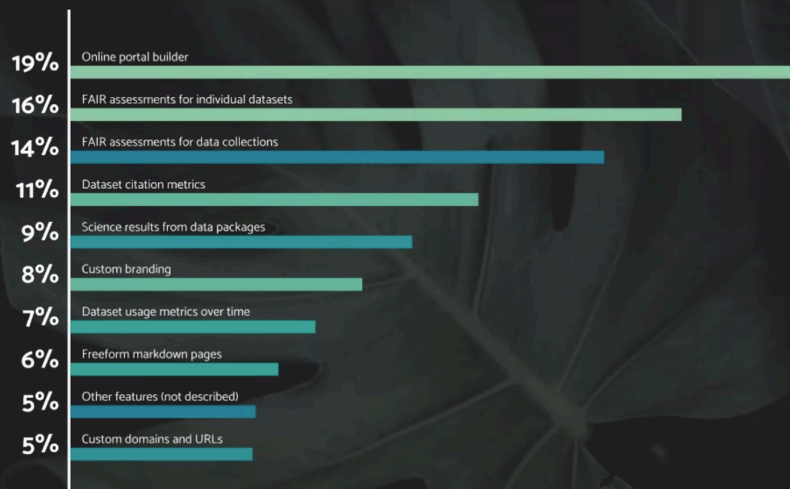
---

#### Highlights from breakouts

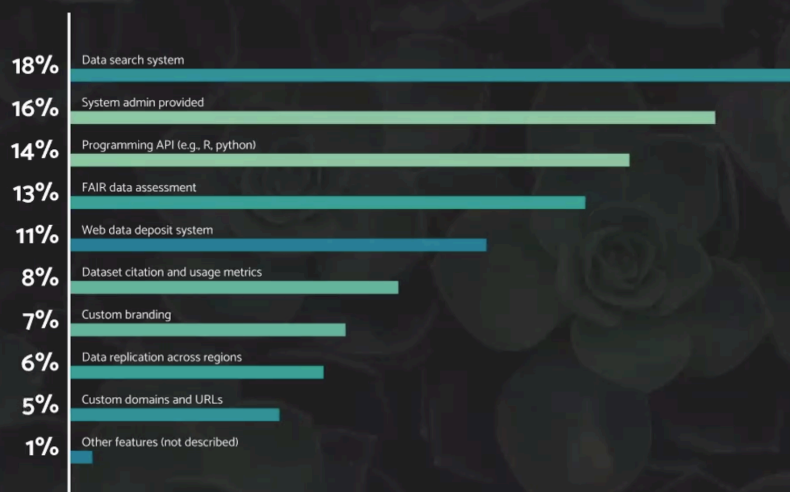
- A:
  - Search capabilities in and out of network, how the search works, and having a better API for that.
  - Also a lot of interest in FAIR checks, and questions about sustainability & preservation when client runs out of funds
- B:
  - Hosted repositories seem good for those who don't want to manage their own tech. Curiosity about storage costs and high volume content. Storage costs - seems like a flat fee, downsides to moving to an annual system, worry about persistence into the future. What type of resources needed on site?
- C:
  - Focus on who are the users for portals, and helping researchers grasp that early on.
  - Future of metadata needs for researchers, standards.
  - Partially automated metadata.
  - NSF funding means big goals - how do portals support related use cases?
  - Different features for different audiences?
- D:
  - What separated DataONE learning services from other learning services
  - Existing training on data portals - where to find
  - Who governs the existing training materials
  - Additional training materials because one-size fits all might not work for all groups

Menti poll

## DataONE Plus services value to you:



## Hosted repository value to you:





## Community Meeting

<https://www.dataone.org/meetings/2020/>

Zoom: 943 0955 6242 Pwd: 3282663

One tap dial in: +16699006833,,94309556242#

[code-of-conduct](#), [@DataONEorg](#)

---

### Q&A:

- What is the open source platform(s) that DataONE is based on?
  - DataONE bases most infrastructure on existing standard - SOLR, POSTGRES
  - MetacatUI, Metacat, R Tools, etc, customizable open source libraries
  - From Chris Jones (DataOne):
    - Mark: To build on what Matt said, the Metacat repository system is just one of a number of software stacks that implement the DataONE Service API. For instance, there is a Python-based repository called GMN (generic Member Node) that implements the API. Other groups have built the API support into their own custom repository systems.
- Is there a link for the fee structure
  - The fee structure is not yet published. We are unable to share publicly until it has been approved by campus.
- Would a university be the group to pay for hosted repo for multiple depts.? Or each dept would be dif hosted repo? Makes a huge dif if \$12K/yr covers multiple depts. that are collaborating. Agree that it's a deal for someone else to be the admin.
  - Proportional to the size of the repo, and how much support will be needed
  - Large organizational membership scaled based on how much participation you expect from DataONE
  - If you come in as a small, medium organization to put up a Hosted Repository across two departments with 20 staff, lower end of numbers
  - Large organization: 70 departments, entire university will access, contract negotiated based on size and number of people that DataONE supports
  - You could have 1 repo for two departments, then use portals to specify data collection and search filters for each department
- How are datasets assigned to a specific team catalog for the DataOne plus? Is there an element in the metadata that is used to assign a dataset to a specific team catalog?
  - Through the UI, build a custom query about what data belongs in a collection.
  - Can use any of the metadata fields to build collections
  - Can also individually list datasets that should be included
  - 
  -
- DataONE formed as a network repositories, and now moving into the space of being/running repos.
  - DataONE has worked hard not to compete with repos that we serve.



Community Meeting

<https://www.dataone.org/meetings/2020/>

Zoom: 943 0955 6242 Pwd: 3282663

One tap dial in: +16699006833,,94309556242#

[code-of-conduct](#), @DataONEorg

---

- We decided to go down this path because there are a lot of institutions that would like to run repos but lack the infrastructure
- We are operating infrastructure so that someone else can own a repo.
  - They may decide to export all their data and move it out
  - Can export it and move it out
- Trying to grow the network/ecosystem
- Will the data be stored at DataONE or another repo?
  - SB and Tenese and amazon cloud, - 3 copies
  - DataONE Is operating the physical infrastructure
  - Not trying to create a repo that would replace any other repository, not a centralized data facility
  - Helping the rest of the community operate repos for their community
- Other repositories may no longer receive funding the future and will have to figure out sustainability on their own
  - We need to advocate for funding agencies to continue supporting repositories and preservation
- Fee based model & sustainability