

**QDR Technical Advisory Board**  
**Progress Report**  
**03.04.2020**

**Report Overview**

This report describes the work of the QDR technical team over the last six months (since 09.04.2019), including progress and adjustments made to the 2019 strategic roadmap, and a new roadmap for 2020.

**Section 1 : Completed and work in-progress**

**Dataverse**

The technical team has been making steady improvements to QDR's installation of Dataverse. The following briefly highlights our work:

- Currently, we are running IQSS release 4.17 of Dataverse on Production. and have 4.19 running in a staging environment (more below).
- In previous reports we have described the development and maintenance of a set of media previews that allow users to view content on Dataverse before downloading. We have now transferred maintenance of the previewers to GDCC. Our interim includes significant efforts to expand previewers for tsv/csv, stata and R, etc.
  - Some community efforts to expand tools
- We are currently working to deploy Make Data Count metrics across QDR data collections.
  - About 1 month out

**Infrastructure**

QDR's system [architecture](#) includes a production, stage, and development environments that are hosted on AWS.

- In November 2019 our team completed an upgrade of operating systems running our servers (Ubuntu 18.04 from 16.04) and a configuration management tool (Chef 13).
  - Relatively smooth transition
- The other major infrastructure work includes migration of our content management system from Drupal 7 to 8.
  - This has proved to be a considerable time investment - in part because of the substantive difference between these versions of Drupal (we are moving from Drupal 7 to Drupal 8.8.2), and in part because we are simultaneously testing and deploying a new theme for QDR's site appearance.
  - We currently have a prototype of Drupal 8 and a new site theme running on a staging environment. QDR staff are working on content editing and quality assurance of this prototype before it is deployed to production.
    - Worked with schema design
  - We anticipate completing this migration and site redesign by end-March

## Section 2: 2019 Roadmap Progress

For 2019 the QDR technical team made progress or completed the following:

1. Support and develop new value-added features for QDR's institutional members.
  - a. Implement Make Data Count metrics and begin to provide customized reports for institutional members [In progress]
  - b. Create and further enhance branded Dataverses (within QDR) for institutional members to take advantage of site redesign [In progress]
    - i. Facilitate use by institutional members
  - c. Improved Social Media Integration (e.g. use open graph and twitter metadata to improve social media post preview) [ Complete ]
2. Improve the capabilities of QDR's dataverse for curating and interacting with qualitative data
  - a. This broad goal includes developing curation tools and workflows embedded in DV repository so as to facilitate better ingest, staging and pre-publication curation activities within Dataverse. This will ideally improve the efficiency of curators work, improve provenance tracking, and help identify (and even reduce) errors in manual tasks that curators need to currently perform outside of DV. Our team is in the early planning and design stages of this work (which we will describe in detail on the TAB call). [Moved to 2020 Roadmap]
  - b. Sebastian Karcher's awesome DVCuratoR  
<https://github.com/QualitativeDataRepository/dvcurator>
    - i. Automates some rote manual curation steps; likely to be re-used and refined in the future
3. Infrastructure Refresh and Improvements
  - a. Upgrade existing versions of tools that run QDR and (nginx, Jenkins, Chef, and OS (Ubuntu), Grafana) [Complete]
  - b. Make better use of Docker / Kubernetes for ease of deployment [Moved to 2020 roadmap]
    - i. Continuing to work at; GDCC have started work on this; expecting to join some of those conversation
  - c. Geographically distributed data storage - beginning with Syracuse and eventually working with a third-party provider, or at least West Coast USA location. [Complete]
  - d. Upgrade Drupal 7 to 8 [In Progress]

## Section 3: 2020 Roadmap

The QDR 2020 roadmap is focused broadly on improving the maintenance of our infrastructure, and deploying new.

Less of a focus on infrastructure; more on features

**Q1 2020**                      Branded Dataverses (with our new styling) [Proof of concept]  
Automated GUI testing (Selenium) - Benefit from and contribute  
to the standardized DV test suite  
*Had used Selenium in the past; hoping to re-use*

**Q2 2020**                      Make Data Count + Automate reporting quarterly (or as time  
period appropriate) for QDR and Institutional Members

Begin LSS-ATI tool work to better support authors producing  
annotations for deposit and publication on QDR  
Implementation not 100% clear  
Looking at frontend developer to help with this – looking for  
pointers – please be in touch.

**Q3 2020**                      Explore Kubernetes / Docker for easing deployment and  
maintenance

Migrate to Drupal 9 (Summer 2020) - This is a very different  
versioned update than in the past - Drupal 9.0.0 is meant to be  
really a security update from 8 (rather than a substantive  
architectural change). Minimal, but valuable for us to upgrade  
(hours to complete).

**Q4 2020**                      Prototype Trusted Remote Storage Agent (TRSA) - and explore  
use cases from Sloan Work (Q4)

Explore support for QDAS archiving (Q4)

Allow links to datasets outside of Harvesting context [for further  
discussion]

Items for broader discussion:

1. Linking to datasets outside of simple harvesting context [Sebastian Karcher]
2. Public reporting / transparency for TT work [Nic Weber]
3. Anonymous peer review in Dataverse [Sebastian Karcher and Nic Weber]

