QDR Qualitative Data Repository

Annual Report 2022

QDR.SYR.EDU

DIRECTORS' NOTE

Since "turning on the lights" in 2014, QDR has become a well-established source of expertise about, and a prolific publisher of, qualitative data in the US and beyond. Our collections have grown in both number and diversity. Initially based largely in political science, many of our new deposits are now from researchers in public health and other disciplines. We also take great pride that the repository is increasingly institutionalized: Our curation procedures are documented in peer-reviewed publications, and our technical platform is thoroughly stable with a steady flow of functional and aesthetic improvements. We continue to meet and exceed the expectations for well-administered data repositories. For example, in 2022 QDR received renewal of its CoreTrustSeal certification.

While increasingly rooted, QDR nevertheless maintains its original entrepreneurial "start-up" spirit, continuing to innovate by developing and introducing new tools (both for curators and for researchers), as well as by helping to advance important debates about transparency in qualitative research and on the ethics of sharing sensitive data.

Community is at the core of QDR's approach, and central to its success as a repository: the communities of qualitative researchers of which QDR is a part and with which we engage daily, the community of open science and open data advocates who share our passion for opening up scientific research, as well the community of data professionals in libraries and repositories, who share our dedication to building scientific infrastructure for the long term. No matter which of these communities you are part of, thank you for your interest in QDR's work. Please be in touch if you have any comments or questions about this report or about QDR's activities.

Colin Elman

Director, Qualitative Data Repository
Professor of Political Science
Director, Center for Qualitative
and Multi-Method Inquiry
Syracuse University

Sebastian Karcher

Associate Director, Qualitative Data Repository
Research Assistant Professor of Political Science
Syracuse University

INSTITUTIONAL DEVELOPMENTS

QDR is a CoreTrustSeal (CTS) certified "Trustworthy Data Repository." This certification means that the repository has met a rigorous set of core requirements established by CoreTrustSeal, an "international, community based, non-governmental, and non-profit organization promoting sustainable and trustworthy data infrastructure." As QDR's initial CoreTrustSeal certification expired in 2021, we prepared and submitted the materials required to apply for re-certification and were awarded re-certification in October 2022. It bears noting that QDR's Associate Director also serves as a reviewer for CTS.

In addition, several prestigious publishing outlets or platforms (e.g., PLOS, Springer/Nature) or began to (F1000Research, Elsevier, eLife) continue to list QDR as a resource and encourage authors to consult with us regarding the management and sharing of qualitative data. We are also mentioned as a useful resource on the web pages of many colleges and universities throughout the United States.

More generally, as detailed in the sections that follow, in 2022 QDR continued to generate and contribute to important conversations, and engage with key institutions, related to research transparency and openness in the social sciences. QDR personnel attended conferences and workshops, were invited to give talks and webinars, and had our work published in important academic and practitioner outlets. QDR is becoming increasingly established and recognized within the open science community, as well as by researchers across the social sciences and beyond.

Two advisory boards offer guidance to QDR, assisting us with evaluating the repository's progress and envisioning its future: a Research Advisory Board (RAB) and a Technical Advisory Board (TAB). Both boards are elected every two years. Typically, some board members opt to remain on the board and others leave, and are replaced by new members suggested by a nominating committee. On January 1, 2022 we welcomed two new members to the RAB and four new members to the TAB. The current configuration of both advisory boards as well as all previous members can be found here. In 2022 we held two meetings with each board, benefiting from the RAB's advice and insights on how to engage with the scholarly community, and from the TAB on how to enhance QDR's technical foundation and profile.

Part of the revenue that supports QDR derives from annual dues paid by academic and other institutions that become institutional members. To build our membership base, QDR personnel engage with data librarians and members of university leadership at institutions across the United States. In the spirit of open science, the dues institutional members pay to QDR support the *curation* of data projects rather than *access* to data. In 2022, we continued our efforts to retain our current institutional members (with more than 95% of institutions renewing their membership) and to recruit new ones. At the end of 2022, QDR had 32 institutional members. With the changes in data sharing policies for the National Institutes of Health (NIH), we have seen a growing interest in QDR membership. We are planning to develop additional benefits

and programs for institutional members to help their researchers, and aid with recruitment and retention.

GRANTS AND FUNDING

QDR relies on two types of external funding: our core general operations grant from the NSF, and a series of project-based grants which support QDR personnel and infrastructure. In 2021 we renewed our core grant, and also received several other major new awards.

Accordingly in 2022, we performed work under five NSF grants: We are in the final months of a grant supporting QDR's general operations and continued development, and started work on the newer core award. Three more specialized grants support our work on Annotation for Transparent Inquiry (ATI), Working with Sensitive Research Data, and the Journal Editors Discussion Interface (JEDI).

We also worked on a grant from the Institute for Museum and Library Science for work on archiving data from Qualitative Data Analysis Software, and subawards from awards by the Alfred P. Sloan foundation (on privacy encoding sensitive data) and the Mellon Foundation (on Indigenous data sovereignty), led by QDR's technical co-directors Nic Weber and Carole Palmer, respectively.

In addition to funding from grants and the annual dues paid by the repository's institutional members, QDR is also supported by deposit fees.

PERSONNEL

In summer 2022, QDR's founding co-director Diana Kapiszewski stepped back from the bulk of her responsibilities at QDR to focus on other academic projects. Diana had an incalculable impact on every aspect of QDR – the repository would not exist without her leadership. While we continue to miss Diana's expertise, enthusiasm and efficiency, QDR has transitioned to a new equilibrium. We are fortunate that apart from Diana's departure, QDR's core personnel otherwise remained stable throughout 2022. QDR's technical infrastructure is developed by two consultants who have been with the repository for a number of years, Dr. Jim Myers and Sebastian Ostrowski. Their work is directed by technical co-director Dr. Nic Weber, in close coordination with QDR's associate director Dr. Sebastian Karcher. With her long and varied experience, QDR's other technical co-director Dr. Carole Palmer brings crucial expertise to the technical team, as well as leadership on a number of substantive topics, including how QDR approaches indigenous data. We also worked closely with To Huynh, who serves as one of the core developers for the "Anno-REP" tool (see "Research and Publications" below).

Dr. Priya Silverstein, who previously served as the Community Manager for the Journal Editors Discussion Interface (JEDI – see "Connections, Outreach, and Promotion"), has started a new position as a Postdoctoral Scholar in the Psychological Science Accelerator. Priya will remain connected to JEDI as its new associate director. We are excited to welcome Julia Bottessini, who has joined JEDI as its new Community Manager.

Finally, during both semesters of 2022 we had teams of outstanding graduate assistants from the Departments of Political Science, and Public Administration and International Affairs, History, and Social Science working with us. The energy and intelligence that Syracuse University students bring to their work at QDR enlivens the repository and strengthen our operations and research.

In an exciting development, QDR has hired postdoc Derek Robey, who is currently finishing his PhD in Sociology at Harvard University. Derek will be joining us as the inaugural Postdoctoral Scholar in Data Ethics and Research Openness in August 2023, funded by the NSF and Syracuse University.

CURATION AND INTERNAL OPERATIONS

In 2022, QDR implemented new substantive facets (search and browse filters) which allow users to find published projects by the types of data they include. The new facet leverages existing functionality of the Dataverse system, on which the QDR catalog is built, and appears alongside other facets for filtering project search results, such as geographic coverage, keyword, and disciplinary subject. Projects can, and typically do, comprise more than one type of data. As can be seen when the facet is used, currently the majority of published data projects contain interview transcripts, with archival documents and coded qualitative data running close behind. Additionally, about a quarter of QDR's holdings are ATI (Annotations for Transparent Inquiry) projects, which received their own dedicated label as well.

During the year, QDR's operations continued to benefit from the focused improvements in the curation process which the curation team made in 2021. With an in-house purpose-built package that assists with common Dataverse-based curation tasks, QDR now automates many of the repeated steps in our typical curation process. This frees up staff time to be spent on the idiosyncratic elements of each deposit that require individualized in-person engagement. The tool, called DV Curator - Python, is tailored for how QDR approaches curation, and so will likely require modifications for other repositories. Nevertheless, its updated code and documentation are publicly available.

In 2022, a new cohort of curators completed their initial self-training in QDR's "human-in-the-loop" approach, learning about the essential curation tasks performed for any project. The training was based on QDR's comprehensive Curation Handbook, augmented by guidance in responsible data handling. (More about the motivations for our approach and some key attributes of the process can be found in Demgenski, Robert, Sebastian Karcher, Dessi Kirilova, and Nic Weber. 2021. "Introducing the Qualitative Data Repository's Curation Handbook." *Journal of eScience Librarianship* 10 (3): 1207. https://doi.org/10.7191/jeslib.2021.1207) Behind the scenes, we inaugurated a new set of status labels which indicate the completion of different curation steps to provide faster internal communication prior to project publication. Lastly, as a result of our 2021 accessibility review of the QDR website, to ensure compliance with the Web Content Accessibility Guidelines (WCAG), we have also incorporated some additional accessibility steps to the data we process as part of curation. While partial and voluntary, these

steps enhance the accessibility of many data file types we publish. Having both the automated and manual steps of the QDR curation process outlined in detailed documentation increases efficiency and consistency, eases staff training, and facilitates the continued evolution and improvement of our processes.

The refinement of operational, policy and technological processes have prepared the repository for ongoing success and stability. Crucially, as two newly formulated sets of federal expectations for trustworthy repositories were published in 2022 – by the NIH and by the White House Office of Science and Technology Policy – QDR was quickly able to document how its set-up already meets them.

DATA RECRUITMENT, ACQUISITION, AND PUBLICATION

RECRUITMENT

QDR actively pursues data recruitment through a variety of ongoing strategies. First, we conduct regular outreach campaigns offering QDR as an appropriate venue for sharing the data from projects of up-and-coming social scientists, recognized as especially meritorious either by national (for example, as NSF DDRIG grantees) or private (for example, Templeton or Wenner-Gren) funding agencies. When researchers are early in their project, QDR offers tailored data management consultations (see more in Outreach section), discussing eventual data sharing options. In addition, we identify exemplary individual works of qualitative social science and reach out to the researchers behind them directly to discuss the possibility of sharing their project's data via QDR. Finally, we strategically identify existing networks and institutions with interest in qualitative social science and research transparency as "pipelines" to data sharing and offer QDR's services to them. We bring an "all-of-QDR" approach to these data recruitment efforts, dedicating a specified weekly time when team members implement these strategies.

DATA ACQUISITION AND PUBLICATION

In the course of 2022, QDR received 109 new deposits that fit our collections development criteria. This was the second year in a row when the number of new deposits has increased significantly compared to the previous year (in 2021, we received 49 new submissions, compared to 37 in 2020.) These deposits comprised a mix of projects recruited by QDR, and "walkins" who found the repository independently, with the latter category comprising an increasing portion of all deposits.

In the course of our initial deposit appraisal, a handful of initiated deposits are deemed not to be a good fit for QDR's scope, most frequently because they do not contain any qualitative data and, more rarely, because they are from a discipline not part of the broadest definition of "social science". In those cases, we try to direct depositors to more appropriate venues. Another portion of deposits QDR has to decline are those that come to us with informed consent

language that does not allow for data sharing after the end of the project and as a result specifies commitments to participants incompatible with data publication in a repository.

QDR published 27 projects in 2022, a number close to the 28 published in 2021. The total number of published data deposits as of December 31, 2021, stood at 147. At the time, over 200 entries were awaiting eventual publication, either because they are in various stages of curation or because the depositors are still collecting data. (QDR encourages interested researchers to initiate a project on the system when still at the planning stage so that we can offer ideas of best practices for data management.) One of our goals for 2023 is to bring fully deposited projects to publication faster than we have been able to in the past.

DATA DEPOSIT COLLABORATIONS

Two types of collaborative initiatives between QDR and external entities interested in data availability started in 2022: one with a prominent Political Science publication and another spurred on by the new NIH Data Management and Sharing Policy.

- We published three supplemental data projects accompanying articles appearing in the journal *Security Studies*, which has a dedicated collection with QDR. We expect an increasing number of deposits from similar empirical articles which use qualitative or multi-method approaches.
- Looking forward to 2023, when the new NIH DMSP comes into effect, QDR worked with two NIH-funded research programs, JCOIN and HEAL, to assist with the archiving and sharing of qualitative data. QDR has qualified as a HEAL-compliant repository and can serve researchers under those programs.

DATA COLLECTION HIGHLIGHT

One of QDR's most recent and active collections, supported by the Palliative Care Research Cooperative (PCRC) with funding from the National Institute of Nursing Research, illustrates the potential for research-group-based compilations clustered around a common research program. Researchers affiliated with institutional member PCRC and motivated by their cooperative's explicit focus on transparency and sharing information with both other scholars and study participants, have provided a steady stream of new data publications over the course of 2021 and 2022 from recent empirical studies in palliative nursing science. The three projects published in 2022 have attracted over 2,000 views, complementing the four which were brought to publication over the course of the previous year, which have attracted over 10,000 views.

REUSE

The top three downloads from QDR (in 2022 and in general) are projects we know are actively used for pedagogical purposes:

- Matima, Rangarirai; Murphy, Katherine; Levitt, Naomi S.; BeLue, Rhonda; Oni, Tolu. 2018. "A qualitative study on the experiences and perspectives of public sector patients in Cape Town in managing the workload of demands of HIV and type 2 diabetes multimorbidity". Qualitative Data Repository. https://doi.org/10.5064/F6CE5MRF. QDR Main Collection. (Used by Alistair Thomlinson since 2019 in a graduate research methods course at Cardiff Metropolitan University, UK)
- Saunders, Elizabeth. 2015. "Data for: "John F. Kennedy," in: Leaders at war: How presidents shape military interventions". Qualitative Data Repository. https://doi.org/10.5064/F68G8HMM. QDR Main Collection. V3. (Used by Robert Adcock since 2016 in an undergraduate historical methods in international relations course at American University, USA)
- O'Neill, Maureen. 2017. "High performance school-age athletes at Australian schools: A study of conflicting demands". Qualitative Data Repository. https://doi.org/10.5064/F6ZP448B. QDR Main Collection. (Used by the depositing researcher and others in NVivo trainings in Australia.)

The next most downloaded data project illustrates an interesting new form of coordinated re-use. It was selected as the source of data to be re-analyzed by a collective of qualitative researchers with different methodological tools. The resulting analyses, as well as an accompanying piece by QDR staff on the role of curation of epistemically responsible re-use, appeared as a special section in the June 2021 issue of *The Qualitative Report*.

- Chukwuma, Adanna. 2017. "A qualitative study of perspectives on postnatal care referrals by Traditional Birth Attendants in Ebonyi State, Nigeria". Qualitative Data Repository. https://doi.org/10.5064/F67H1GGS

CONNECTIONS, OUTREACH, AND PROMOTION

QDR's outreach aims to connect with several distinct audiences relevant to social and health science data: researchers, other data professionals, as well as institutional stakeholders with an interest in data and open science, including journals, institutional review boards (IRBs), disciplinary associations, and funders publishing the work of and supporting the research of the two primary communities of interest. In all our engagement with these individuals and institutions, we also attempt to draw connections and facilitate conversations amongst them, driven by the belief that a holistic approach leads to more effective and sustainable data practices.

SOCIAL SCIENCE, HEALTH SCIENCE AND COGNATE RESEARCHERS

QDR's impact extends across the social and health sciences and beyond. Always intended as a repository that would serve a broad spectrum of scholarly disciplines involved in the study of humans, in all their individual behaviors and social relationships, QDR's most recent outreach

events – many by invitation – demonstrate that our reach now extends far beyond the repository's initial site in Political Science.

In June 2022, QDR was invited to join a panel on ethical and human subjects considerations for data sharing from the repository angle, as part of an ICS3-organized workshop "Advancing Data Sharing for Implementation Science in Cancer Control." This presentation came on the heels of the 2021 SAGE/NVivo webinar on "Theory and Practice of Qualitative Health Research" where QDR staff joined depositors from public health and demography studies, who have worked with us. In both of those events, addressed to the health science community, we presented general concepts of research transparency and how they relate (and do not relate) to qualitative work, as well as considerations around confidentiality and informed consent and different strategies for responsible data sharing.

As part of the annual Institute for Qualitative and Multi-method Research (IQMR), in June 2022, QDR personnel provided 31 data management plan (DMP) consultations to participants. This number was lower than the 45 sessions we were able to offer as part of IQMR 2021, due to the virtual format that year, but was still the highest number for the five years during which we have offered this option at in-person meetings.

Also in June 2022, QDR staff offered an invited presentation on sharing qualitative data in Linguistics, as part of a Cambridge University Press-sponsored symposium "Open Scholarship in Applied Linguistics" during which we also offered a live demonstration of an ATI project from the field of Linguistics published on QDR.

In September 2022, QDR personnel taught a short course for political scientists as part of the annual meeting of the American Political Science Association (APSA) on "Managing and Sharing Qualitative Data." This short course was designed to equip participants with a range of strategies for effectively managing qualitative data, including general good practices for data management, documenting analysis conducted in qualitative data analysis software (QDAS) and employing Annotation for Transparent Inquiry (ATI). Hands-on exercises allowed participants to practice basic data management tasks in the context of their own projects. The short course particularly emphasized writing DMPs (a long-standing requirement for proposals to the National Science Foundation (NSF)), for research involving qualitative and multi-method data. We have offered this short course for six years and it remains a well-subscribed offering at APSA. QDR also hosted a booth at the conference.

In November, the Collaborative Institutional Training Initiative (CITI) invited QDR staff to deliver a webinar on "The New NIH Data Management and Sharing Policy for Social, Behavioral and Educational Researchers and IRBs." Over 250 people registered to attend.

In December, QDR staff presented a paper on "IRB Guidance on Data Sharing in the Social Sciences: Evidence of a Shift" at the annual conference of PRIM&R (Public Responsibility in Medicine & Research, the flagship professional organization of IRB personnel) documenting an observable growth in the number and level of detail of mentions of data-sharing possibilities in IRB guidance at key US research universities.

DATA PROFESSIONALS

Based on the voluntary accessibility self-audit QDR conducted the previous year and described above, in 2022, personnel involved in that effort presented their insights and some initial recommendations for good practices in this unexplored area in digital curation at the 2022 annual meetings of the RDAP (Research Data Access & Preservation) Association, the International Digital Curation Conference (IDCC), and Code4Lib organizations. QDR staff also delivered the invited keynote talk on the same topic at the STEM Library Association meeting. The IDCC presentation was published in the *International Journal of Digital Curation* with former QDR graduate assistant Theresa Anderson as a co-author.

In 2022, the Data-Facing Track of the Campus Research Computing Consortium (CaRCC) hosted a call dedicated to "Curating, Archiving and Sharing Qualitative Data" with QDR's associate director. For a different session, QDR staff co-presented on "Teaching Research Data Management to Students" alongside data librarians. Videos and notes from both virtual events can be found online.

By invitation, we also introduced QDR both as a repository venue and as a resource for secondary data of particular use for college students to groups of social science librarians via online presentations at the Politics, Policy, & International Relations (PPIRS) Section and the Anthropology and Sociology Section of the Association of College and Research Libraries (ACRL), a division of the American Library Association.

GENERAL

QDR regularly makes use of a blog embedded on its own website, where during 2022 we published a description on how QDR meets the new federal US requirements for data repositories published by NIH and the White House. We also were excited to publish two guest posts on our blog: In a post by Jessica Nina Lester (Indiana University), Noah Goodman (Education Development Center's Center for Children & Technology), and Michelle O'Reilly (University of Leicester), they describe a recent special issue they edited in *The Qualitative Report* in which multiple groups of qualitative methodologists re-analyzed a single data project from QDR. A second post on re-use, by Marina Mohd Hamdan (West Virginia University), describes how she reanalyzed data by Corey Shdaimah, shared on QDR, for her Master's thesis in Sociology.

Additionally, we use broader social media channels to engage with our communities of interest as well as the academic community at large via a Twitter and YouTube presence. At the end of 2022 we had 1,515 Twitter followers (280 of whom were new for the year). Given growing concerns about developments on the Twitter platform, we have significantly reduced our use of Twitter since November 2022 and are principally using Fediverse with a newly created Mastodon account.

JEDI

The Journal Editors' Discussion Interface (JEDI) is a project hosted by QDR on behalf of Data-PASS. It "provides the community of social science journal editors a forum in which to ask and answer questions, share information and expertise, and build a fund of collective knowledge." Questions cover all aspects of editorial work, but focus in particular on open science practices. The JEDI mailing list is an active Google group with more than 450 editor members. Insights from these discussions are compiled in a large and growing list of resources on the JEDI Website. In May 2022, JEDI held a well-attended virtual meeting, including keynotes by Simine Vazire and Tony Ross Hellauer, as well as breakout groups on topics relevant to editorship.

PARTNERS

The Institute for Qualitative and Multi-method Research (IQMR) is a two-week methods training program held annually at Syracuse University. IQMR promotes the teaching and application of qualitative research methods and their combination with complementary analytic techniques in political science and cognate disciplines. Due to the COVID-19 pandemic, IQMR was held on-line in 2021, but was back to an in-person format in 2022. As in previous years, QDR personnel held individual DMP consultations for all interested IQMR attendees, helping them to strengthen their plans and practices relating to data management.

QDR continued its active engagement with the Data Preservation Alliance for the Social Sciences (Data-PASS), a consortium of organizations (mainly data repositories) dedicated to preserving data generated through and used in social science research. In particular, with QDR as the lead organization, Data-PASS continued its work in establishing the Journal Editors Discussion Interface described above.

WEB PRESENCE

QDR's [website](#) is at the center of our activities. The site provides information for scholars who wish to deposit, use, or re-use data, instructors of methods and other courses, and students at all education levels – in short, anyone interested in qualitative data and research. At the end of 2022, QDR had 4,540 registered users, 997 of whom registered during that calendar year. Registration is not necessary for viewing most information on QDR, however, so our website reaches a much larger audience than just our registered users. About 22,000 unique visitors browsed QDR in 2022 in about 57,500 sessions, a roughly 10 percent increase over the previous calendar year. The following five pages were the most visited pages on our site beyond the homepage. They demonstrate that researchers visit QDR to learn about the institution, but also for information on topics specific to qualitative data, such as de-identification and data management:

- [Deposit Process](#)
- [Managing Data](#)
- [About the Qualitative Data Repository](#)
- [Types of Qualitative Data](#)
- [De-Identification](#)

RESEARCH AND PUBLICATIONS

As a complement to our activities curating, preserving, and publishing qualitative research data, QDR also engages in research on a range of topics related to data sharing and research openness more broadly. Substantial progress was made on each QDR research project, as detailed below.

Annotation for Transparent Inquiry (ATI) – ATI is an approach to transparency in qualitative and multi-method research that enables scholars to annotate passages in the text of a published work, elucidating how the underlying data were generated and analyzed, and potentially linking to the data themselves. ATI made significant advances in 2022: a workshop at Syracuse University's Lubin House in New York City brought together a group of researchers working on law and social sciences who used ATI for their work and discussed its impact. Their feedback also informed further development of the [Anno-REP tool](#), which is now publicly available and has been used to facilitate the publication of a number of ATI projects, including, for the first time, a book using ATI. All empirical chapters of *[Constitution Makers on Constitution Making](#)*, edited by Tom Ginsburg and Sumit Bisarya, include ATI annotations accessible directly on Cambridge University Press's Core platform, as well as as standalone documents on QDR.

Qualitative Data Analysis Software (QDAS) – with the support of funding from the Institute for Museum and Library Studies, this project studies how users of qualitative data analysis software (QDAS) create, store, and share their coding work. We fielded a survey to better understand how different qualitative researchers use QDAS tools as well as their attitudes towards sharing such materials. We [pre-registered the descriptive analysis](#) and some initial hypotheses about

data sharing. Simultaneously, QDR added a basic previewer for REFI-QDA files that allows users to look through codes and files in a project deposited in QDR prior to downloading it. REFI-QDA is a relatively new, open exchange format for QDAS projects that has the potential to revolutionize how such projects are shared.

Qualitative Verification – QDR is working to develop systematic procedures for verifying qualitative research, and the qualitative components of multi-method research. "Verification" entails a scholar different from the author of a piece of scholarship analyzing the same data employed in the scholarship, using the same methods, to see if similar results obtain. We drafted an initial set of "Qualitative Verification" guidelines and are using them to verify 5 systematically (for the most part) selected political science articles that use process tracing.

Data Services for Indigenous Data – As part of work funded by the Mellon foundation, QDR is a contributing to a two-year project on Data Services for Indigenous Scholarship and Sovereignty (DSISS). QDR's associate director joined a workshop at the University of Washington and presented on how repository infrastructure helps to protect sensitive data. We also co-proposed a paper to be presented at the 2023 IASSIST Annual meeting with other project participants.

Sensitive Data – QDR has two projects focused on sensitive data. In connection with the "Working with Sensitive Research Data" (WSRD) project, we continued our work under the "Optimizing Openness in Human Participant Research" grant from the NSF's Office of Advanced Cyberinfrastructure, which focuses on connecting Institutional Review Boards (IRBs) and data professionals on issues related to sharing sensitive data. Specifically, we analyzed the guidance regarding sharing human subjects data offered by IRBs at research universities across the U.S., which formed the empirical presentation given at PRIM&R as listed in Outreach. (Other data from the project, generated through interviewing and conducting focus groups with IRB personnel, has been published since 2021: Kirilova, Dessislava; Kapiszewski, Diana. 2021. "Optimizing Openness in Human Participants Research: Harmonizing Standards for Consent Agreements and Data Management Plans to Empower the Reuse of Sensitive Scientific Data". Qualitative Data Repository. https://doi.org/10.5064/F6RQA7AQ.)

Under a "CUSE/G2G" seed grant from Syracuse University, we began exploratory research on offering scholars guidance on keeping sensitive research data safe and secure and sharing them ethically, and facilitating scholarly communication with various stakeholders in the academic ecosystem. We also continued to develop a series of modules on managing and sharing human participants data commissioned by the CITI research, ethics, and compliance training program.

In connection with the "Privacy Encodings for Sensitive Data" (PESD) project, supported by the Sloan Foundation, we conducted more than 50 interviews across four domains of qualitative research and are currently preparing those interviews for publication. We published the first preprint (currently under review for publication) based on these interviews. We also built an R library that facilitates the search of awarded grants across a large number of public and private

funders in the US, to help us to contextualize the collection of qualitative data in the funding landscape and submitted it to the R-OpenSci collection of libraries after passing peer review.

TECHNICAL INNOVATIONS

Over the last year, QDR has continued to improve its repository platform to better support qualitative data, enable the work of its curators, and improve the general performance and accessibility of its holdings.

- QDR now has automated tools to collect customized reports on usage statistics. We demonstrated a version of these on a Dataverse community call (and used the tool to collect some of the information presented in this report).
- We also significantly improved our DVCurator tool, now available with a graphical user interface and written in Python, to automate many of our curation steps and tracking. The tool was presented during the Dataverse community meeting.
- To further enhance security and accountability to stakeholders, QDR published a detailed data breach policy that outlines responses to various data breach scenarios.

As part of our **commitment to free/libre and open-source software** (FLOSS), we have contributed work in each of these areas back to the core Dataverse codebase, maintained at the Institute for Quantitative Social Science at Harvard University, so that users in any of the more than 70 Dataverse repositories worldwide can benefit from QDR's work.

UPDATES, TESTING, OUTAGES, AND SECURITY

Beyond these visible enhancements, QDR's technical team works diligently to keep all relevant **software updated and secure**. QDR is commonly using the latest version of the Dataverse repository software within 2 weeks of its release. As the most significant security improvements during 2022, QDR is now encrypting all data and back-ups at rest using industry best practices (data have always been encrypted in transfer). To ensure the reliability and availability of back-ups, we performed a series of emergency "drills," testing our ability to restore our services in different scenarios. We also started using StorJ as an alternate, fully encrypted and distributed S3 storage alternative to Amazon, particularly for large projects.

We closely monitor security developments and announcements concerning all software components used in QDR. We also undergo regular automated security testing by the University of Texas's "Dorkbot" service.

QDR IN NUMBERS – 2022

977

Number of new QDR users who registered in 2022. We're delighted to welcome all of you to our community!

Number of countries in which authors of published deposits are located

26

41,913

Total number of views across QDR in 2022

Total number of published data projects in [QDR's catalog](#) at the end of 2022

147

3,183

Total number of downloads across QDR in 2022