QDR
QUALITATIVE DATA REPOSITORY

2017

Access

A Note from QDR's Leadership

Welcome to the inaugural edition of *QDR Access!* This and future issues will provide an update on QDR's priorities, activities, initiatives, and plans for the future. We look forward to sharing recent developments with our users, our collaborators, and others invested in the careful management, preservation, use, and re-use of data generated through qualitative and multi-method social science research.

As calls for increased data access and research transparency have resonated across the social sciences, scholars are increasingly grappling with issues that are at the core of QDR's mission. The central challenges involve sharing qualitative data and making qualitative and multi-method research more transparent in systematic ways while attending to logistical, legal and ethical concerns. Addressing these challenges has never been more vital. QDR is proud to contribute to the generation of solutions by creating infrastructure and tools to help scholars make their qualitative and multi-method research more open.

Review of Activities and Progress to Date

Since QDR's launch a clear set of priorities has guided our work, and we have made significant progress with regard to each. Readers can learn more about our initiatives in connection with each dimension in subsequent pages of *QDR Access*.

We have made good progress on our most mission-critical task, developing the repository's holdings. To date QDR has published, curated, and processed for long-term preservation 16 of the 21 pilot projects we commissioned at the repository's launch, and seven non-commissioned data projects. We are actively engaged with many other scholars about depositing their data. We have also developed deep expertise in the unique challenges posed by ingesting, curating, and publishing qualitative data, and penned and promulgated practices and guidance to help scholars manage and share such data.

We have worked to integrate QDR into and engaged with the social science community with the goals of helping to standardize the sharing of qualitative data, and contributing to the construction of a new consensus among scholars, publishers and editors, disciplinary and sub-field associations, and domain repositories. We have hosted or co-organized several workshops on a range of related topics and QDR personnel have published an extensive series of related articles and chapters. QDR has also sought to inform social scientists about the merit and mechanics of managing qualitative data through short courses and other pedagogical efforts.

On the technical front, we have developed QDR's core user-facing and repository functions, and built a robust system for curating data projects. We are in the final stages of adopting Dataverse as QDR's repository platform, transitioning our infrastructure to Amazon Web Services (AWS), and redesigning our web site architecture. We have made good progress

integrating QDR into the data management community in the United States and contributing to the overall development of domestic data infrastructure. In addition, we have begun to generate a full suite of clear policies so that all of QDR's operations are efficiently managed and appropriately overseen.

With regard to institutional development, we have built solid governance structures, both within Syracuse University (QDR's host institution) and externally. QDR's two advisory boards, comprising deeply respected members of the social science and data science communities, offer critical advice and guidance on all aspects of QDR's work. In addition, we are fortunate to have a top-notch, highly skilled staff. Finally, we have worked to ensure QDR's long-term sustainability by leveraging its data and guidance materials, services, and technology to attract funding and to lay the groundwork for institutional memberships.

Preview of Future Priorities and Initiatives

In the months and years ahead, QDR will continue to make progress on each of the dimensions just mentioned. In the near-term, we will prioritize a set of key tasks.

First and most importantly, we will continue to grow QDR's holdings. Building QDR's holdings entails persevering in our efforts to address difficult intellectual, sociological, and technical challenges. We have implemented a multi-faceted data-recruitment strategy that leverages QDR's activities to encourage deposit.

One very promising development involves a new approach to transparency in qualitative research, annotation for transparent inquiry (ATI). Deploying ATI and encouraging its use will involve close coordination with the scholarly publishing community, in collaboration with our technology partner Hypothesis. Moreover, since reuse is a natural complement to deposit, we will continue to pursue a vigorous strategy to encourage the download of QDR data.

With regard to technical priorities, a central objective will be completing and building on the previously mentioned AWS and Dataverse transitions. We will also work to improve public facing technical and other documentation, with the goal of obtaining Data Seal of Approval certification. We will enhance our metadata schema, and explore implementing a DataTags system. We will also seek to improve the user experience and build a meaningful metrics dashboard for tracking user activity. Finally, we will continue to develop QDR's capacity to offer differential access to the data projects we publish, whether they were deposited for secondary analysis, are serving as a supplement to published work, or were developed and shared for pedagogical purposes.

Making progress on these key tasks will result in a more robust QDR and help to solidify its position as an important part of data infrastructure in the United States. We invite your comments and feedback on QDR's work, your participation in our activities, your support, and, of course, *your data*!

Institutional Foundations

In this inaugural edition of *QDR Access* we include a section offering some key institutional background to the founding and subsequent development of the repository. We provide information on what motivated the creation of QDR and describe its mission and value propositions, discuss funding and sustainability, and outline personnel and governance.

QDR is the first U.S. domain repository dedicated to publishing research data arising from qualitative and multi-method social science inquiry. The repository is hosted by the Center for Qualitative and Multi-Method Inquiry (CQMI), which is affiliated with the Moynihan Institute of Global Affairs, a unit of the Maxwell School of Citizenship and Public Affairs at Syracuse University. The Maxwell School is a national leader in multi-disciplinary, practical, problem-driven, social science. Launched in 2014, CQMI offers training, programming, and events related to qualitative data, qualitative and multi-method research, and the processes of making research transparent. QDR has benefited significantly from Syracuse University's strong commitment to establishing the repository as a durable entity servicing social science disciplines in the United States and beyond.

Four beliefs underpin QDR's mission:

- data that can be shared and reused should be;
- evidence-based claims should be made transparently;
- teaching is enriched by the use of well-documented data; and
- rigorous social science requires common understandings of its research methods.

History, Mission, and Value Propositions

QDR's launch in was designed to fill a gap in American social science infrastructure: the absence of a domain repository dedicated to qualitative data. This lacuna posed a serious impediment to including qualitative research in the openness revolution. Sharing and re-using qualitative data require developing new understandings and purpose-built technological capabilities. Moreover, in order for data sharing and re-use to be conceived of as pressing (or even relevant) concerns for researchers operating in qualitative social science traditions, they need to be framed as inherently valuable enterprises with logics that are *consistent* with those traditions. Developing this technology and helping to shift the epistemic and sociological consensus on data sharing among scholars who engage in qualitative inquiry motivated QDR's founding and continue to inspire our work today.

The repository's overarching goals are to catalyze progress toward making qualitative data sharing a customary practice in the social sciences; to broaden the access of scholars around the world to social science data; to make qualitative and multi-method research more rigorous and credible, and to contribute to the enhancement of qualitative research methods; and to encourage intellectual exchange, stimulate the development of epistemic communities, and serve as a platform for global networks and partnerships among scholars and practitioners.

To achieve those ends, QDR ingests, curates, stores, preserves, publishes, and enables the download of digital data arising from qualitative and multi-method research in the social sciences. Ancillary efforts support these core activities and stem from QDR's commitment to addressing the needs and interests of qualitative scholars. The repository develops and disseminates methodologically informed guidance for managing, sharing, citing, and reusing qualitative data, and seeks to contribute to the generation of common standards for doing so. We also train scholars in these practices, and more generally seek to inculcate in the qualitative social science community the importance of sharing the fruits of research, making scholarship transparent, and reusing each other's research data. In short, we encourage and support researchers and help them to achieve openness.

Four beliefs underpin QDR's mission: data that can be shared and reused should be; evidence-based claims should be made transparently; teaching is enriched by the use of well-documented data; and rigorous social science requires common understandings of its research methods.

QDR provides value to a variety of stakeholders in the social science and data management communities including social science researchers, academic publishers and journal editors, data repositories and data librarians, and funding organizations. The repository also provides value to the Maxwell School at Syracuse University.

QDR offers personalized curation and long-term archiving (with access controls as needed) for *social science researchers* who want to share qualitative data (for secondary use, to make their work more transparent, and/or to facilitate teaching). QDR also publishes a growing catalogue of searchable and browseable collections of well-documented qualitative data. For *academic publishers and editors*, QDR provides solutions that allow the books and articles they publish and edit to be appropriately, efficiently, and economically connected to associated qualitative data and materials.

For *repositories* that share QDR's goals of archiving and preserving social science data and promoting data sharing, QDR offers productive partnership and a willingness to work jointly to develop common solutions to challenges pertaining to the management of qualitative data. For *data librarians* who advise researchers who work with qualitative data, QDR provides expertise and written guidance on qualitative data management.

For *funding organizations* invested in demonstrating best scholarly practices and maximizing the impact of the research they fund, QDR offers advice and tools to aid grantees in effectively demonstrating the rigor and power of their research and sharing its fruits for use by others. With respect to *funders that generously support* QDR, the repository strives to use grants frugally, efficiently, and to maximum possible effect, as it seeks to address funders' goals and aspirations.

To QDR's host institution, *Syracuse University*, QDR offers visibility and multiple services benefiting faculty and students. Nested within the broader set of institutions comprising the infrastructure of qualitative and multi-method research located at Syracuse, QDR forms part of a multi-faceted system supporting these critical forms of inquiry.

Funding and Sustainability

The funding QDR receives from external institutions powers its operations, reflects trust and confidence in the repository, and serves as a valuable marker, augmenting QDR's credibility and legitimacy.

From the outset, QDR has benefited from the generous support of the National Science Foundation (NSF), which has provided the majority of the institution's core funding. We are especially indebted to Brian Humes, director of the Political Science program in the Directorate for Social, Behavioral, and Economic Sciences. More broadly, the NSF's support for data management and the building of data infrastructure across the natural and social sciences has been seminal in starting and sustaining the movement towards greater openness.

Earlier this year, QDR received funding from the Robert Wood Johnson Foundation to support the continued development and evaluation of a new approach to transparency in qualitative and multi-method research, annotation for transparent inquiry (ATI). Readers can learn more about this exciting and powerful transparency technique in the Technical Infrastructure section of this report.

QDR plans to introduce institutional memberships for which academic and other organizations will be eligible to apply. Researchers, faculty, students and staff at member institutions will have access to enhanced services, training, and tools provided by QDR. More news on this initiative will appear in subsequent editions of *QDR Access*.

Personnel and Governance

QDR sits at the intersection of the social science and data management communities, and the repository's personnel and governance structures are rooted in both constituencies.

QDR fields an outstanding core staff who are trained to be attentive to the special challenges that managing and sharing qualitative data pose, and who are committed to addressing the multiple concerns that qualitative social scientists have in connection with data sharing. Most members of the repository's "Content Team" are located at Syracuse University. Since 2015, QDR has benefited enormously from a partnership with the Information School at the University of Washington, where the repository's technical leadership is located. Finally, two consultants form our technical team, both located remote from Syracuse and Seattle. While coordinating such a distributed team introduces some complications, these are outweighed by the benefits of having such a talented group of contributors, each of whom brings key expertise to the project.

OUR TEAM

Colin Elman—Director (Syracuse University)

Diana Kapiszewski—Deputy Director (Georgetown University)

Sebastian Karcher—Associate Director (Syracuse University)

Dessislava Kirilova—Curation Specialist (Syracuse University)

Carole Palmer—Technical Director (University of Washington)

Nicholas Weber—Technical Director (University of Washington)

Sebastian Ostrowski—Systems Administrator

Alex Ivanov—Drupal Developer

Christiane Pagé—Associate Director, Center for Qualitative and Multi-Method Inquiry (Syracuse University)

Sarah-Anna Hogan—Office Coordinator, Center for Qualitative and Multi-Method Inquiry (Syracuse University)

With regard to institutional governance, QDR has established two distinguished advisory groups: a Research Advisory Board and a Technical Advisory Board. Each board meets semi-annually to be apprised of the repository's progress and to offer advice and guidance. QDR has benefited significantly from the support and input offered by both boards.

RESEARCH ADVISORY BOARD

John Gerring (University of Texas at Austin)
James Goldgeier (American University)
Marcus Kreuzer (Villanova University)
Sean M. Lynn-Jones (Harvard University)
James Mahoney (Northwestern University)
David Mainwaring (Cambridge University Press)
Andrew Moravcsik (Princeton University)
John Owen (Chair) (University of Virginia)
Amy Pienta (University of Michigan)
Elizabeth Saunders (George Washington University)
Anne-Marie Slaughter (New America Foundation, Princeton University)
Marc Trachtenberg (UCLA)
Lisa Wedeen (University of Chicago)
Jennifer Widner (Princeton University)

TECHNICAL ADVISORY BOARD

Micah Altman (Massachusetts Institute of Technology)
Thu-Mai Christian (University of North Carolina at Chapel Hill)
Louise Corti (UK Data Service)
Mercè Crosas (Chair) (Harvard University)
Timothy DiLauro (Johns Hopkins University)
Gary King (Harvard University)
Margaret C. Levenstein (University of Michigan)
James D. Myers (University of Michigan)

As QDR's holdings and activities multiply, we will continue to develop our core staff and governance structures, broadening QDR's reach into our two key constituencies and enhancing the repository's profile and the expertise on which we draw.
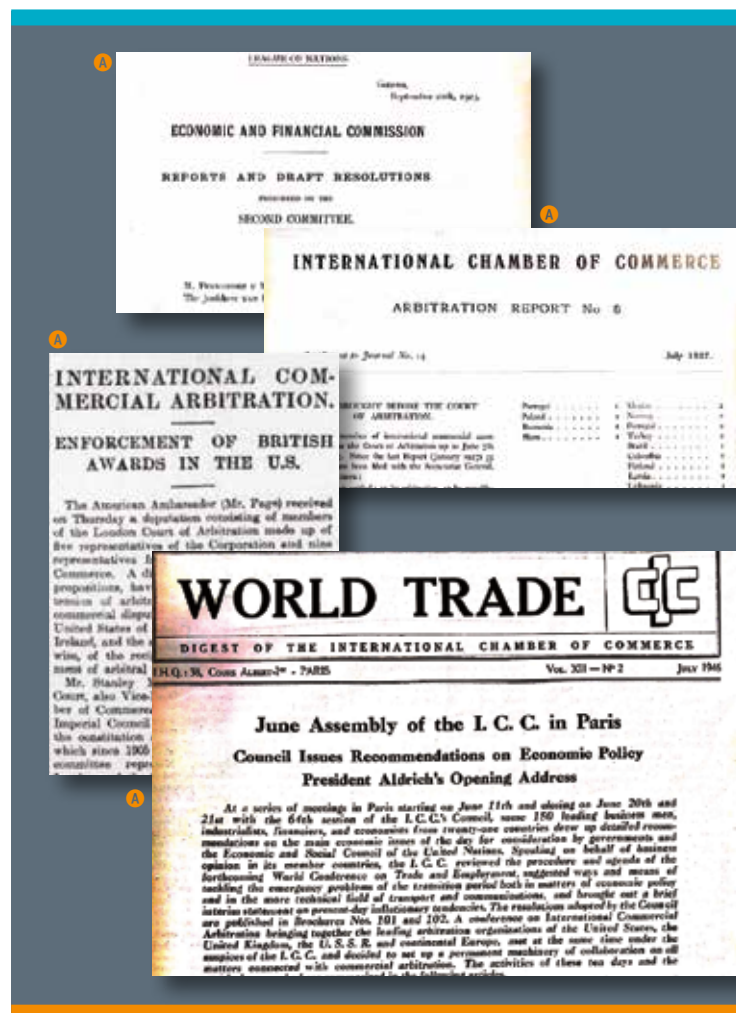
Collections

Consistent with its relatively recent launch, QDR hosts a small but steadily growing collection of data projects covering a wide range of historical periods, topics, and geographical regions. For instance, we host data projects relating to educational policy in 19th century Peru, transnational commercial disputes, Russia's foreign policy leading up to World War I, and attitudes towards post-natal care in Nigeria.

The 23 data projects that scholars have deposited with QDR to date are of two types. Seven are "active citation compilations." Based on foundational work by Andrew Moravcsik (Princeton University), active citation compilations accompany a research publication; they provide extracts of sources that underlie the publication, and include annotations explaining how those sources support claims made and conclusions reached in the publication. (See the "Technical Infrastructure" section of this report for exciting news about annotation for transparent inquiry [ATI], an approach to making qualitative research transparent that QDR is developing, building on active citation.) The remaining 16 data projects comprise 843 data files freely accessible to registered QDR users, and several hundred additional files the access to which is restricted due to privacy or copyright concerns.

The diverse data hosted by QDR provide insights into the many forms that qualitative social science data can take. For instance, QDR holds de-identified interview transcripts as well as transcripts from focus groups. Other data come in the form of reports or narratives, e.g., a compilation of brief descriptions of the attitudes towards democracy and dictatorship held by important political actors in 20th century Latin America. QDR also hosts tables and spreadsheets, for instance, a list of newspaper articles relevant to judicial decisions. Yet another project comprises campaign videos from Latin American presidential elections.
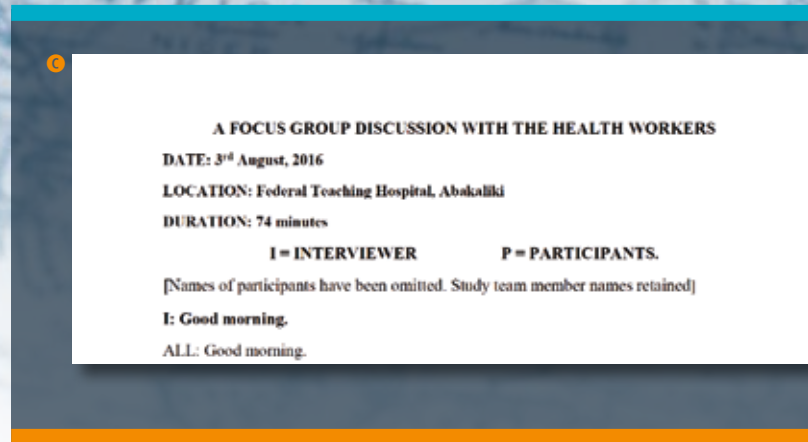
A closer look at three data projects illustrates the richness and diversity of QDR data deposits. The data in a project deposited by political scientist Thomas Hale (University of Oxford) entitled "Between Interests and Law" (http://doi.org/b465) relate to research on institutions for

resolving international commercial disputes. The data comprise notes from interviews with policy makers and other experts from five countries, scans of archival documents, key documents from both national legislatures and international bodies, as well as historical newspaper coverage. Together, they help readers to better understand a research project that spans four continents and five centuries. Hale's project exemplifies the diversity of qualitative data used in contemporary social science.

In his data project "The Rise of Ethnic Politics in Latin America" (http://doi.org/b4m4), political scientist Raúl Madrid (University of Texas at Austin) provides notes from interviews with 47 political leaders and experts from Bolivia, Ecuador, Guatemala, and Peru; an in-depth description of the data; and documentation of the informed consent language that he used when interviewing. The data underpin his book of the same name (Cambridge University Press, 2012) and several journal articles. Madrid's work seeks to explain why some indigenous-based parties in Latin America fail while others succeed. The project provides candid insights into the worldviews of key participants in the processes Madrid studies. Importantly, the project highlights that the data drawn from interviews conducted by social scientists are not always sensitive. Madrid's interlocutors all have extensive media

Ⓐ Hale, Thomas. (2016) "Between Interests and Law: The Politics of Transnational Commercial Disputes." Data Collection, QDR:10069. Syracuse, NY: Qualitative Data Repository [distributor]. http://doi.org/b465.

Ⓑ Boas, Taylor. (2015) "Presidential Campaign Advertising in Chile, Brazil, and Peru." Data Collection, QDR:10053. Syracuse, NY: Qualitative Data Repository [distributor]. http://doi.org/b5zd

Ⓒ Chukwuma, Adanna, Chinyere Mbachu, Jessica Cohen, Margaret McConnell and Thomas Bossert. (2017) Data for: "A Qualitative Study of Perspectives on Postnatal Care Referrals by Traditional Birth Attendants in Ebonyi State, Nigeria." Data Collection, QDR:10078. Syracuse, NY: Qualitative Data Repository [distributor]. http://doi.org/b4nr

experience and agreed for their words to be attributed to them, allowing their publication without de-identification.

Finally, scholar of public health Adanna Chukwuma (Harvard University) shared data underpinning a specific manuscript that is currently under review. The piece examines the conditions under which traditional birth attendants refer clients to health services for postnatal care. "Data for: A Qualitative Study of Perspectives on Postnatal Care Referrals by Traditional Birth Attendants in Ebonyi State, Niger" (http://doi.org/b4nr) includes de-identified transcripts from three focus groups (in this instance, de-identification only required removing the names of the focus group participants). The collection also includes the detailed guides Chukwuma used to structure each discussion, documentation of informed consent language, and IRB approval. Chukwuma's data project demonstrates how providing qualitative data in tandem with a research publication can facilitate understanding and evaluation. By sharing all of the focus group discussions on which the scholarship is based (rather than just providing brief transcript excerpts in the immediate text of the article), Chukwuma and her co-authors allow readers to check whether their conclusions are warranted by those discussion as a whole. The project also highlights QDR's growing holdings from disciplines outside political science.

If you are a qualitative or multi-method researcher in the social sciences, we look forward to talking with you about sharing your research data through QDR.

Workshops

Fulfilling QDR's central mission—facilitating the sharing, archiving, publication, and reuse of qualitative data—requires addressing crucial questions about how those processes can be carried out legally and ethically. QDR organizes thematic workshops with invited experts to develop strategies to address these challenges. To date, workshop topics have included how copyright law, human participant ethics and institutional review board (IRB) regulations, and qualitative data analysis software interact with data sharing. The main goal for each event was to derive a set of consensus guidelines that can assist QDR with our work *and* serve as an input to ongoing conversations about standards and practices for sharing data and making research transparent.

"Sharing Data in the Social Sciences: Copyright Challenges," co-sponsored by the Social Science Research Council (SSRC), took place in April 2015. The workshop gathered copyright lawyers, university librarians and archivists familiar with copyright issues, and prominent advocates of research transparency in the social sciences. The conversation was framed by the central tension between greater access to scholarly products and the legal mandate to respect intellectual property. Attendees debated issues such as fair use, international aspects of copyright in the globally connected academic community, how technology can facilitate access to qualitative source materials (text, video, images, etc.), and how the current impulse toward data sharing may impact the scholarly production of various communities of practice.

Conversations began from the consensus premise that the legal risks that sharing some types of qualitative data might present can only be managed, not eliminated altogether. Thus, workshop participants sought to imagine a realistic roadmap for the academic community for designing better advice and clearer policies on sharing (and managing the risk of sharing) qualitative data that are under copyright. Subsequently, we incorporated much of this information into the guidance materials we provide to users and into our internal policies. Several participants generously agreed to continue sharing with QDR their expertise and recommendations with regard to copyright issues.

QDR held another workshop, "Protecting People and Sharing Data," in May 2016. The workshop focused on the ethical and legal concerns that can impact sharing social science data generated through interactions with human participants. The workshop's main task was to generate practical solutions for both individual researchers and repositories/archives, drawing on real-world case studies and considering general best practices for data management. The diverse set of attendees included scholars with deep expertise on IRBs, legal experts who have advised other organizations, researchers working with different types of human participants in various difficult contexts, and seasoned data management professionals.

The group recognized the benefits that accrue from broadening access to data. They also acknowledged how ongoing changes in technology, evolving expectations of data sharing, and revisions to the Federal Common Rule complicate the task of generating generalized guidelines for

sharing data. Attendees sympathized with QDR's position of advocating for the sharing of qualitative data *and* accepting that access to some data needs to be restricted. They likewise supported our efforts to build tools that facilitate data management planning and might allow for some qualitative data to be shared even in difficult cases. The biggest workshop takeaway was that many dilemmas can be avoided or made less acute if research is conducted with eventual data sharing in mind, and if scholars use mitigating strategies that are already available. In the course of the discussion, QDR identified the critical role that IRBs can play in the process of communicating these strategies to researchers. Subsequent editions of *QDR Access* will report on our follow-on initiatives involving engaging members of the broader IRB community in a discussion of differential access to data.

In October 2016 we convened a workshop on "Computer-Assisted Qualitative Data Analysis (CAQDAS) Projects and Digital Repositories' Best Practices." CAQDAS software—which allows scholars to manage, link, tag, annotate, produce memos about, analyze, and visualize raw qualitative data—presents new challenges for sharing qualitative data. The workshop considered the difficulties researchers face in sharing CAQDAS output and underlying data, and explored technical solutions, for example an interchange-and-preservation format. The timing of the workshop allowed QDR to leverage ongoing efforts to create a strong and flexible XML archiving format, linking across two promising projects: the QuDEx format prototype from the UK Data Archive, and recent efforts by KWALON (the Netherlands Association for Qualitative Research) to create a minimal-components standard. Full implementation and adoption of these formats are still in the future. Nonetheless, the meeting represented a valuable opportunity to share strategies and lay the groundwork for generating technical guidelines and practical templates for CAQDAS archiving—materials that could be used as the basis for building software in the future.

The knowledge QDR gained through these workshops has been extremely helpful as we have developed guidance on these important questions, and we are indebted to all of the workshop participants. However, much work still needs to be done on these and other topics, and we look forward to continued engagement with scholars active in qualitative research and data management.

Skill–Building

Instantiating data sharing and re-use requires intellectual and sociological shifts. Traditional scholarly practices are deeply embedded and their stickiness presents a major obstacle to change. As such, providing helpful information to social scientists—on managing, sharing, and reusing qualitative data; on making qualitative and multi-method research transparent; and on the benefits of doing so—is a critical aspect of QDR's mission. Our long-term objective is to help make these practices customary. Through engaging in teaching and generating guidance materials, we believe we have begun to make a difference.

Each year QDR personnel offer a module on "Managing and Sharing Qualitative Data and Making Qualitative Research Transparent" at the Institute for Qualitative and Multi-Method Research held annually at Syracuse University. We also offer a short course on the same topic at the American Political Science Association (APSA) annual meeting. These modules and short courses, which we will continue to offer into the future, have already been attended by dozens of social scientists. In each, we demonstrate procedures for and the value of effective data management throughout the research lifecycle; highlight the benefits of sharing data; consider perceived barriers to data sharing and describe techniques to overcome them;

and discuss how qualitative researchers can make their work more transparent and the benefits of doing so. QDR has also begun to offer webinars. The first, "Managing Qualitative Data Safely and Securely" (based in part on a course previously co-taught by personnel from QDR and the UK Data Service), was presented in November 2016. Around 50 participants—including a number of non-U.S. researchers—attended the presentation.

QDR also offers an ever-expanding set of "Guidance and Resources" (http://bit.do/dqv2B), as well as links to a regularly updated list of relevant conferences, workshops, and training offered by other institutions. Two elements of these materials merit highlighting. First, in line with our view that shared data can be used to enhance instruction in research methods, we offer guidance on "teaching with data." Readers are also encouraged to view an example of a "pedagogical data project" (which can be used in teaching) deposited by political scientists Jeb Barnes (University of Southern California) and Nick Weller (University of California, Riverside) (http://doi.org/b48p). Second, QDR has assembled customizable teaching materials on managing and sharing qualitative social science data. After piloting their use at Syracuse University, we shared them via email with faculty who teach qualitative research methods courses in 90 top Ph.D. programs in political science, sociology, and anthropology, and also posted them to the QDR web site (http://bit.do/dq3bP). While primarily designed to be used for a single week of a research design or related methods seminar, the material could also serve as the basis for a "Data Management" session in professional development workshops for graduate students. We are also very willing to connect with individual Ph.D. candidates to offer them hands-on data management assistance.

QDR also aspires to engage with and help other stakeholders and constituencies within the social science community. We have a particular interest in interacting productively with the editorial teams at social science journals to assist them as they incorporate data-related processes—e.g., citing and verifying data and making them accessible—into their workflows. One effort in this regard was our contribution to an August 2016 workshop organized by the Data Preservation Alliance for the Social Sciences (Data-PASS, please see the "Outreach" section of this report for details). The journal editors and others who attended were provided with a summary of recent developments in data citation, management, and sharing, and learned about how these improvements could be integrated into journal workflows. The next workshop in this continuing series will be held in connection with the Dataverse Community Meeting in Boston in June 2017.

While QDR aims to reach beyond Syracuse University, we are also active at our host institution. The Center for Qualitative and Multi-Method Inquiry (CQMI), which houses QDR, holds events related to data management and sharing, and research transparency. Individual QDR staff also actively engage with the university community. In October 2016, for instance, QDR staff discussed ongoing debates on, and the role that digital repositories can play in, research transparency in the social sciences with the Moynihan Research Workshop, a group of junior faculty from departments across the Maxwell School.

Further, as noted elsewhere in this report, since QDR's inception we have benefited from the efforts of terrific Graduate Assistants (GAs) and Research Assistants (RAs) who work with QDR for an academic year (AY), most hailing from various programs in the Maxwell School. Since 2014, our GAs have included David Arceneaux (AY 2013-14), Colleen Burton (AY 2015-16), and Arianna Galluzzo (AY

2016-17). Our RAs have included Lucy Peijing Xu (AY 2013-14), Evan Kindle (AY 2014-15), Nicholas Southwick (AY 2015-16), and Jinsol Park (AY 2016-17). We take seriously our mission of training these young scholars, and seek to involve them in our work in multiple ways, in particular with data curation. For example, Colleen co-authored a note on research cycles with QDR Director Colin Elman, and Colleen and Arianna both attended a Social Science Research Council (SSRC) workshop in March 2017 on "Curating Knowledge under Digital Conditions." Colleen had this to say about her experience working with QDR: "It was interesting to see how fields other than political science handle questions about research transparency and philosophies of data sharing... Overall, my year at QDR was a great learning experience and I am incorporating what I have learned—especially regarding planning ahead for data management and sharing—as I move forward on my dissertation and other research projects." We look forward to integrating more students into our curation team over time!

> "It was interesting to see how fields other than political science handle questions about research transparency and philosophies of data sharing... Overall, my year at QDR was a great learning experience and I am incorporating what I have learned—especially regarding planning ahead for data management and sharing—as I move forward on my dissertation and other research projects."
>
> — COLLEEN BURTON (AY 2015–16)

Technical Infrastructure

QDR's technical team is led by Carole Palmer and Nic Weber, both faculty in the Information School at the University of Washington. The technical team is responsible for the overall functioning of the repository. Its members work to build new features for QDR, increase its functionality, ensure that QDR safely preserves data, and keep its systems running.

QDR's website—qdr.org—provides a wealth of information on qualitative data management and sharing for QDR users and others. Connected to that site is our data repository system, which stores data and makes them available for searching and browsing. In the background are another set of services that monitor the site's security and performance, create multiple back-ups of our data holdings and of the site, and connect us to other organizations, such as the California Digital Library's EZID service, through which QDR registers its Digital Object Identifiers (DOIs).

Looking forward, by the end of 2017 we expect to complete a dual technological transition that will improve the user experience and make QDR's infrastructure more scalable and reliable, as well as help us to obtain Data Seal of Approval certification.

First, QDR is adopting Dataverse repository software (dataverse.org), open source software developed at Harvard's Institute for Quantitative Social Science. Due to its intuitive user interface and active development community, Dataverse is rapidly gaining a diverse worldwide user base including such institutions as the Odum Institute at the University of North Carolina, the Dutch National Archives, and the Beijing University data services. Transitioning to Dataverse will immediately improve QDR's data deposit process, curation capacity, and access management. Dataverse's attractive user interface features powerful search and facetted browsing options as well as built-in metrics. Following the transition, QDR will actively contribute to the Dataverse software community, focusing on innovations for qualitative data such as tools to handle text and image files.

Because Dataverse software is open source, these improvements will be available for use by all Dataverse installations, improving the curation of qualitative data far beyond QDR.

Second, QDR is migrating its infrastructure to Amazon Web Services (AWS). This transfer will provide us with a more flexible set-up. For instance, running separate environments for different services will augment security and performance. The migration will also allow QDR to quickly scale its services. In contrast with local hardware solutions, cloud services such as AWS allow us to requisition additional storage space, bandwidth, or even entire server instances within minutes when the need arises. Before deciding on the use of AWS, QDR conducted an in-depth security analysis based on the Cloud Security Alliance's Consensus Assessment questionnaire, and

	A	B	C	D	E	F
	R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.					
1	QDR Inventory for compliance with DSA Common Requirements: https://docs.google.com/document/d/1_DPwSA5P6LpK9Q34BhxJmX8So2GKL7eSLa-G-z5JvVg/					
2	Requirement	Specific Question	QDR Fulfills	Relevant URL	Comments	
13	R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.					
14		How does the repository comply with applicable disciplinary norms?				
15		Does the repository request confirmation that data collection or creation was carried out in accordance with legal and ethical criteria prevailing in the data producer's geographical location or discipline (e.g., Ethical Review Committee/Institutional Review Board or Data Protection legislation)?	In Progress		Yes we do, but we don't have a formal policy written out	
16		Are special procedures applied to manage data with disclosure risk?	Yes		Special Deposit and Download agreements	
17		Are data with disclosure risk stored appropriately to limit access?	Uncertain		We think so -- certainly when DV is up	
18		Are data with disclosure risk distributed under appropriate conditions?	Yes		Special Download agreements	
19		Are procedures in place to review disclosure risk in data, and to take the necessary steps to either anonymize files or to provide access in a secure way?	Yes		Data Curation	
20		Are staff trained in the management of data with disclosure risk?	Yes		We teach this, though may need to include GAs	

QDR's Data Seal of Approval progress matrix

team members presented the findings at the International Data Curation Conference in February 2017. We also consulted with various peer institutions that are already running some or all of their services on AWS to assess potential concerns, costs, and benefits.

An additional priority for 2017 is completing QDR's certification as a trustworthy digital repository. While there are several standards available for such certification, the Data Seal of Approval (DSA) is one of the most widely used. Gaining DSA certification entails demonstrating compliance with 16 core requirements, each specified in an assessment questionnaire that is completed by the applying repository and evaluated by a peer. This self- and peer-assessment process provides us with useful checks on our policies and proce-

dures, ensuring that they conform to internationally recognized best practices. DSA certification will also send a strong signal, confirming QDR's status as a trustworthy digital repository. QDR already fulfills most DSA requirements and is actively progressing on the few that are outstanding.

Part of QDR's mission is providing services that address the needs of scholars who engage in research involving qualitative data. As standards and requirements for making social science research more transparent are emerging, qualitative scholars need strategies and tools for compliance. To assist them, QDR is developing a new approach to transparency in qualitative research, annotation for transparent inquiry (ATI) (http://bit.do/dovnh). Many qualitative researchers invoke qualitative data sources—interview transcripts,

historical documents, pictures—throughout their manuscripts to underpin descriptive and causal inferences or interpretations. Making such research transparent requires linking those sources to the relevant passages in the manuscript, and offering authors an opportunity to include analytic notes detailing the connections between data sources, and claims made and conclusions drawn in the manuscript.

ATI leverages open web annotations, which recently achieved status as an official web standard, and open source software provided by a technology company called Hypothesis, to achieve these linkages. Text on a publisher's web site is annotated using Hypothesis and linked to data sources hosted by QDR. You can see an early example of ATI accompanying a working paper by Sam Handlin (University of Utah) published by the University of Notre Dame's Kellogg Institute for International Studies here: http://bit.do/qdr-ati-handlin. QDR has received funding from the Robert Wood Johnson Foundation to evaluate ATI over the next two years, and to plan its further development. We are especially keen to work with researchers who are interested in annotating their scholarly work. If you have a publication that you would like to enliven and empower in this way, we look forward to hearing from you.

An annotation from: Handlin, Samuel. (2015) Data for: "The Politics of Polarization: Governance and Party System Change in Latin America, 1990–2010." Annotation for Transparent Inquiry, QDR:10065. Syracuse, NY: Qualitative Data Repository [distributor] http://bit.do/qdr–ati–handlin."

Outreach

QDR is committed to being an active member of, and to engaging with a broad range of stakeholders from, both the data management and social science communities. QDR works with data librarians, personnel from other repositories, social science and data science researchers (faculty and graduate students), representatives of public and private foundations, journal publishers and editors, and members of regional and international data alliances, to promote knowledge about the management, sharing, use, and reuse of qualitative data. These activities are instantiated in a variety of formats and settings.

Our publication page (http://bit.do/doNoj) offers a window on one form that our outreach takes: contributing to ongoing methodological and technological debates through publication. QDR personnel have published on various topics including research transparency, data management, and data citation. Our work has appeared in journals such as *Perspectives on Politics*, *PS: Political Science & Politics*, *Security Studies*, and *IFLA Journal*, as well as in the *Comparative Politics Newsletter*, *Qualitative & Multi-Method Research*, and *D-Lib Magazine*. Further, QDR staff regularly present their research at international conferences in both social and information science.

Another form QDR outreach takes is institutional relationships. We have been especially impressed by the generosity and warm encouragement of other repositories. From QDR's very early days, personnel from UK Data, the Inter-university Consortium for Political and Social Research (ICPSR), and the Odum Institute have offered us their advice and shared their expertise. We were also delighted to be invited to join the Data Preservation Alliance for the Social Sciences (Data-PASS), a voluntary organization of domain repositories. Members of Data-PASS cooperatively identify, appraise, acquire, catalog, and preserve data used for social science research.

We have been an enthusiastic participant in joint activities Data-PASS has organized. For instance, in 2016, Data-PASS sponsored an event, "Clarifying Data Citation and Sharing—Some Practical Information for Journals," timed to coincide with the American Political Science Association (APSA) annual meeting. The event aimed to help journal editors consider issues relevant to making data available in tandem with published articles to promote research transparency. Journal editors' willingness at the workshop to share their expertise inspired a follow-on initiative, the launching of the Journal Editors Discussion Interface (JEDI) (https://dpjedi.org/). Data-PASS has now established a workshop series, "Developing and Implementing Data Policies: Conversations Between Journals and Data Repositories," and will sponsor more events on related topics for journal editors. One, entitled "Evolving Practices for Data Management and Sharing," will take place in tandem with the June 2017 Dataverse Community Meeting.

QDR also engages in outreach through organizing its own events, co-organizing events, and participating in events held by other organizations. Readers can learn about the workshop series that QDR sponsors in the "Workshops" section of this report. Further, early in 2017 QDR co-organized "Love Your Data Week", a social media event and campaign to generate tips, resources and best practices for data management, preservation, reuse,

and library/repository research-data services. QDR's Associate Director Sebastian Karcher represents QDR on the organizing committee, which includes specialists from multiple academic institutions. QDR also coordinated with JSTOR to organize a data rescue event in New York, which brought together data specialists to secure crucial data from the US Department of Agriculture.

QDR personnel frequently collaborate with scholars from other institutions. For example, Sebastian is a member of the Data Citation Implementation Project (DCIP) project of FORCE11, and co-authored "A Data Citation Roadmap for Scholarly Data Repositories" (http://doi.org/b4zc) with Martin Fenner (DataCite), Mercè Crosas (Harvard University), and other DCIP members. QDR Director Colin Elman and Arthur Lupia (University of Michigan) co-edited a journal symposium on "Openness in Political Science: Data Access and Research Transparency" (http://doi.org/b42w). In service to the International Association for Social Science Information Services & Technology (IASSIST) community, Sebastian, together with Louise Corti and Libby Bishop (UK Data Service) offered a workshop on preparing qualitative data for sharing and reuse at this year's IASSIST conference.

Involvement in the governance of other data management organizations and cognate institutions is yet another form of QDR outreach. Colin and Sebastian are co-chairs (with Columbia University's Kerstin Lehnert and Jamie Shiers of the European Organization for Nuclear Research [CERN]) of the Research Data Alliance (RDA) Domain Repositories Interest Group. Colin has served on the ICPSR Governing Council since June 2015. He currently chairs the Summer Program Sub-Committee and additionally serves on the Collection Development Committee. Colin is also a member of the Center for Open Science's Transparency and Openness Promotion (TOP) Guidelines Coordinating Committee, a member of the Curating for Reproducibility (CURE) Advisory Group, and was previously a member of the RDA Organizational Advisory Board. Sebastian is

a member of the Services and Technology Steering Group of DataCite. We are fortunate to have several leaders from the data management community serving on QDR's advisory boards. These two-way governance interactions help QDR cultivate collaborative relationships that strengthen our ability to provide state of the art informational resources.

All of these forms of outreach inform the guidance and resources that QDR develops for its user base and the qualitative social science community more generally on topics such as data management, citing data, and teaching with data, described in the "Skill-Building" section of this report.

QDR also has a well-developed social media strategy. We publish a blog (http://bit.do/dovxu) and write regularly about news and events related to qualitative data and their management. QDR also maintains a presence on Twitter (@qdrepository) and reaches out to depositors through quarterly email updates. Recent communications focused on QDR events, annotation, and our partnership with open source annotation non-profit innovator Hypothesis (https://hypothes.is/).

Finally, QDR contributes to open source initiatives relevant for data citation and accessibility. For example, we are currently adopting the Zotero open source reference management software to support QDR affiliated projects, including the Methods Coordination Project (MCP) (http://bit.do/dq4CX) a bibliographic resource covering a range of data collection techniques and analytic methods employed in social inquiry. MCP is co-directed by Colin and John Gerring (University of Texas).

We consider outreach to be an important part of our mandate as a member of the data management community and as a valuable component of data infrastructure in the United States. We welcome contact from organizations engaged in initiatives to which QDR could contribute.

QDR
QUALITATIVE DATA REPOSITORY

QDR ingests, curates, stores, preserves, publishes, and enables the download of digital data arising from qualitative and multi-method research in the social sciences. The repository develops and disseminates methodologically informed guidance for managing, sharing, citing, and reusing qualitative data, and seeks to contribute to the generation of common standards for doing so. QDR's overarching goals are to catalyze progress toward making qualitative data sharing a customary practice in the social sciences; to broaden the access of scholars around the world to social science data; to make qualitative and multi-method research more rigorous and credible, and to contribute to the enhancement of qualitative research methods; and to encourage intellectual exchange, stimulate the development of epistemic communities, and serve as a platform for global networks and partnerships among scholars and practitioners.

QDR
QUALITATIVE DATA
REPOSITORY