IJDC | *Research Paper*

Facilitating Access to Restricted Data: Operationalizing Trust in Data Users

Allison R. B. Tyler
School of Information
University of Michigan

Abstract

The decision to allow users access to restricted and protected data is based on the development of trust in the user by data repositories. In this article, I propose a model of the process of trust development at restricted data repositories, a model which emphasizes the increasing levels of trust dependent on prior interactions between repositories and users. I find that repositories develop trust in their users through the interactions of four dimensions – promissory, experience, competence, and goodwill – that consider distinct types of researcher expertise and the role of a researcher's reputation in the trust process. However, the processes used by repositories to determine a level of trust corresponding to data access are inconsistent and do not support the sharing of trusted users between repositories to maximize efficient yet secure access to restricted research data. I highlight the role of a researcher's reputation as an important factor in trust development and trust transference, and discuss the implications of modelling the restricted data access process as a process of trust development.

Received 07 December 2018 ~ *Revision received* 30 October 2019 ~ *Accepted* 30 October 2019

Correspondence should be addressed to Allison R. B. Tyler, University of Michigan School of Information, 4322 North Quad, 105 S. State St., Ann Arbor, MI, 48108. Email: arbtyler@umich.edu

An earlier version of this paper was presented at the 13th International Digital Curation Conference.

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: http://www.ijdc.net/

Copyright rests with the authors. This work is released under a Creative Commons Attribution Licence, version 4.0. For details please see https://creativecommons.org/licenses/by/4.0/

Introduction

Trust is a central feature of human interaction and forms the basis on which social, political, and organizational exchanges occur. In organizations, for example, differing levels of access granted to visitors and employees alike reflects the level of trust an organization places in individuals (Levenstein, Tyler and Davidson Bleckman, 2018). This trust is an outcome of the confluence of value sharing, identity verification, conflict of interest mitigation, and social capital. When the organization is a digital data repository, a central challenge that has emerged is the development of the trusted digital repository as a secure and trusted place for research data to be curated and disseminated, as identified through the Trusted Repository Audit and Certification and the Data Seal of Approval, among others. Other researchers have focused on identifying the trustworthiness of the digital objects held in data archives and repositories (Donaldson, 2016; Rieh, 2002). However, these do not reflect another important trust development process at repositories with restricted data: the trust in data users by data repositories.

Many data repositories curate both publicly accessible and restricted data. I define restricted data as any data that, due to concerns over individual privacy, security, or commercial interest, have access restrictions in place. Researchers wanting to make use of these data may be required to request access not just for multiple datasets, but to multiple repositories, with no ability to transfer the access one repository has granted to another. Each repository has its own process and criteria for data access request approval, including how identities are verified and incorporated into trusted digital access identities which are affiliated with specific data security and access requirements. While there are legitimate concerns over security and confidentiality that engender these procedures, for researchers who are considered trusted users of restricted data, the inability to transfer these trusted identity credentials between datasets and between data repositories hinders research efforts that would maximize the usability of research data.

In this paper, I evaluate the processes by which repositories with restricted data determine who, and to what degree, to trust with restricted research data. I ask the following question: what model best explains the process of trust development by data repositories of data users?

The purpose of this study is to develop the theoretical foundation for a digital researcher credential that transfers the trust one repository has placed in a data user to another repository. To understand this process, I developed an adaptation of the Boersma, Buckley, and Ghauri (2003) model of transactional trust development to evaluate the processes of trust development at restricted data repositories.

Background

Four concepts underlie this research: what is the importance of data reuse, how 'trust' is currently understood, how trust is operationalized through identity, and why understanding trusted digital identity creation is important for expanding access and reuse of research data.

Data Reuse and Access

Research data have value beyond their original collection purpose, including cost savings, increasing the potential analytical value of data, analysis replication, and validation (Corti, 2007; Manhas et al., 2015). The last decade has witnessed an increasing support among researchers about sharing their data with other researchers, though concerns remain about how the data may be used and who is reusing the data (Tenopir et al., 2015). Bolstered by internet-based data archives and repositories, and funder, journal publisher, and institutional requirements to deposit research data in data repositories, more research data have joined government and organization-produced administrative data in being available for secondary analysis and reuse (Holdren, 2013; National Institutes of Health, 2003; Productivity Commission, 2017; SpringerNature, 2017). These data are valuable resources for researchers interested in comparative studies and historical research, in addition to reducing the burden placed on research respondents and maximizing the utility of the reused data. Social science data, such as interviews, survey responses, and observation field notes, are valuable resources for secondary analysis and reuse on their own or in conjunction with new research that seeks to answer new questions. However, concerns exist about the reuse of these data, especially when these data contain sensitive or protected data which are restricted in use. Access to these data, if allowed, is often mediated through legislative and institutional policies identifying both who is authorised to access these restricted data, and by what means (Eschenfelder and Johnson, 2014; European Parliament, 2016). The process and criteria by which repositories make these decisions about who they allow access to which data differ significantly between repositories, as we found in this study, and challenge the goal of increasing reuse and collaborative, comparative studies across multiple institutions.

Trust

Trust and trustworthiness are the basis of human interaction: trust is the act of putting faith in or taking a risk about a person or object, while trustworthiness is a characteristic or attribute of the trustee (Akter, D'Ambra and Ray, 2011; Kelton, Fleischmann and Wallace, 2008; Yoon, 2014). Wynne (1992) disagrees that trustworthiness is an intrinsic attribute of any person, arguing instead that trust and trustworthiness are instead relational based upon social relationships between people. In both perspectives, the development of trust implies a transaction between the trustee and the person or institution trusting them, with rights and responsibilities inherent in that transaction as well as consequences when that trust is violated. Trust is not a perfect process, and trusted individuals throughout history have broken faith with those who trusted them, including at the highest levels of government. Despite this, trust development remains an important component of the decision-making process in restricted data repositories.

Identity as Trust

In recent years, discussions of digital identity have focused primarily on social media identity. A common interpretation of digital identity is the result of all the interactions a user has online, from blog posts and tweets to the user networks created through online interactions (Ertzscheid, 2016; Sullivan, 2012). While social media access credentials can be access credentials for many other internet resources, from online shopping to the Inter-university Consortium for Political and Social Research, these are not the types of

digital access identities I focus on in this analysis. In repositories, information provided through data access requests is used to authenticate the user to the system via a digital identity and is used to authorize behaviors. This authentication is accomplished through some level of identity verification, including background checks and verification of the user-provided information and affiliations. Once that is completed, a trusted digital identity is created that reflects the amount of trust a repository is willing to place in a user based on the intersection of various dimensions of trust, and this intersection thus determines how researchers can access the data they need.

Importance of Trusted Digital Identities

As an increasingly common research practice, data reuse is a way to maximize the productivity of data gathered through public and private funding. The question of how to maximize data reuse when the data require special permissions for access and are held at multiple institutions underscores the goals of this project. At present, an access credential from one repository does not transfer to a different setting, which can complicate and delay research while awaiting access at each individual repository or for each individual dataset. The goal of this research into how repositories are developing trusted access identities for their users is to build just such a model, utilizing a framework of trust based on the Boersma et al.'s (2003) model, to do so. Restricted data access credentials reflect the willingness of repositories to trust users to behave correctly with the data they are entrusted with, as evidenced by the levels of restriction associated with these data (Levenstein et al., 2018). The results of this research will expand those trusted relationships that individual repositories develop out to a larger consortium of repositories such that users can leverage the trust they have engendered elsewhere for a common purpose. This will provide a better understanding of the nature of the trust relationships between repositories and users, and allow for a more refined model of trust as it relates to repositories and identities of access.

Theoretical Framework

Boersma et al. (2003) present a process model for trust development in international joint venture (IJV) negotiations. The definition of trust that informs this model is the 'expectation that a party can be relied on to keep to agreements (promissory), will perform its role competently (competence) and that the party will behave honourably even where no explicit promises or performance guarantees have been made (goodwill)' (Boersma et al., 2003). This definition presents trust as the interaction between negotiating parties along three dimensions: promissory-based, competence-based, and goodwill-based. This process is recursive; at each stage in the negotiation process, different dimensions are prominent, and the output of each stage serves as the input in the next (see Figure 1).

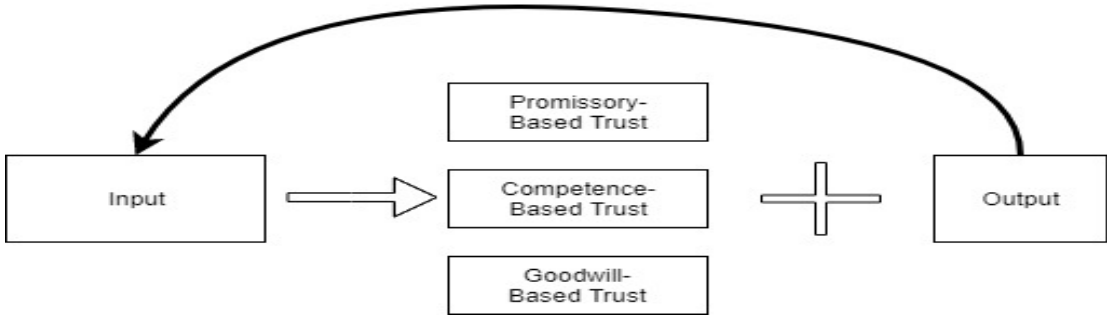

Figure 1. The Boersma, Buckley, and Ghauri model of International Joint Venture negotiation trust development.

My findings indicate that the decision-making that restricted data repositories perform when evaluating restricted data access requests can be modelled as a similar process of trust development. In restricted data repositories, I argue that with each iteration of the process, the amount of trust that is the output of the process increases, provided that the trusted researcher does not provide cause to lose that trust. What I propose in this paper is a revised model that adds an additional dimension of competence and redefines the dimensions to fit the restricted data repository processes. This model, discussed below, reflects the different interactions and considerations operationalized through the data access process. This new model of trust development (Figure 2) is embedded within that larger process (Figure 3).

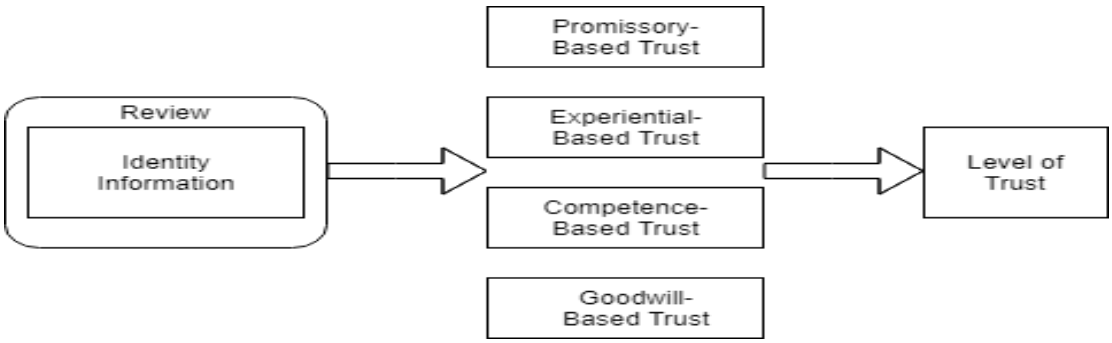

Figure 2. The Model of Trust Development for Restricted Data Repositories.

Methods

I conducted a qualitative study based on interviews and document analysis to understand how repository staff understand trust development. Twenty-three restricted data repositories from five countries were selected using a mix of dichotomous and typical case sampling (Schensul and LeCompte, 2013) based on their size, affiliation, type of data, and their curation and preservation of restricted data. Over 355 pieces of documentation were collected from repository websites and interviewees, including all available information about data security levels, data access request forms, data use agreements, and any other miscellaneous documentation related to how users are able or prohibited from accessing data. Nine interviews were conducted with repository staff.

The interviews were conducted with an aim to understanding actual policy implementation and supplement the existing documentation. They enabled triangulation of findings between the documented policies and interview responses and provided insight into how repository staff understood the role of trust (Levenstein, 2019). The documentation and the interviews were analysed using the qualitative data analysis software NVivo. The code set used was created through an iterative process of grounded coding from the documentation (Saldaña, 2016; Strauss and Corbin, 1990).

Findings

To address the question of what the process of trust development is within restricted data repositories and begin to understand the types of trust that are developed, I structured my findings around four categories that emerged from the documentation and interviews that reflect how repository staff described their review processes: Identity, Training, Reputation, and Project. These categories are not distinct; as seen in Table 1, there is overlap between components of the categories, and certain categories are more prevalent than others. Separating the required components into categories is important for identifying inconsistencies and conflicts between repositories that would inhibit a transfer of trust, as well as commonalities that will inform standardized policies. What I show is that there are inconsistencies in these requirements across repositories. These inconsistencies are not only in how many repositories require the components of each category, but in how the categories themselves are prioritized and defined.

Table 1. The most commonly required components of data access requests, and the number of repositories with explicitly stated requirements.

Identity		Training		Reputation		Project	
Name	18	Data privacy and confidentiality	4	Institutional affiliation	18	Statement of purpose	15
Institutional affiliation	18	Responsible data use	2	Institutional role	14	Signed data use agreement	13
Contact information	13	Information security	2	Training completion	6	Data security plan	8
Country of residence	10	Disclosure control	2	CV	3	IRB/Ethics review	6
Supervisor information	7	CIPSEA	1			Project timeline	5
CV	3	Title 13/26	1			Justification for data	5

Identity

The required Identity components are the most standardized across repositories in data access request forms, data use agreements, or account applications. Eighteen of the 23 repositories require some form of contact information and institutional affiliation. Name, email address, institution name, and institutional role are the fields most commonly verified by repository staff. This is done through searching institutional websites for the given name, and matching the name and email address to the role the

website indicates. A common delay in the approval process is incomplete or inaccurate information at this stage.

> 'The other thing you do need to establish in your registration is your institutional affiliation. So, if you submitted something saying, 'I'm at the University of Michigan,' but you put in your Gmail and it's impossible for anyone to go to the University of Michigan and find you anywhere, you will also not get access to the data' (SF006_1).

Training

I identified two different types of training and experience of importance to repositories. The first reflects specific academic skills that researchers develop during their university career, as evidenced by their degree-attainment and current institutional role. Fourteen repositories require that researchers provide their institutional role in data access requests, and seven explicitly require students, including doctoral students working on dissertations, to provide this information about their supervisor(s). No repositories explicitly state accommodations for commercial researchers who may have the experience but not the advanced degree or faculty position that act as proxies. This reflects the concerns held by several interviewees about principal investigator (PI) or research team member experience with restricted data even if no requirements are made. It did not matter if their respective institutions currently required evidence; if they did require evidence of academic experience, that was 'good'; if they did not, in the representative's opinion, they should. The expected researcher behaviour with data depends on the level of experience the researcher has at the outset.

> 'That's part of the reason for making a distinction between the user types. In order to be a graduate student or an honor student or a staff member… there will often be institutional affiliation, as well. … There's a set of expectations based upon your affiliation and the previous experience you've got' (SF001_1).

The second reflects the non-academic certifications mandated for researchers separate from their specific academic qualifications (e.g., responsible conduct of research training as a requirement for IRB approval, or confidentiality and legal requirements training for security clearance approval). There was an unexpected finding here: despite nearly universal agreement among the interviewees that knowing that a researcher had been trained in restricted data management and security and understood the importance of confidentiality, only six of repositories require any sort of extra training prior to accessing restricted data, and there is little consistency between those institutions as to what training is required. Interviewees were asked about their repository policies for data security and data management training, and while few actually require it, the interviewees from institutions that did not still agreed that it is important to train researchers on how and why they should protect these data. One interviewee, when asked about repository training requirements, said

> 'No, but I wish we did. … We're working now on basic training for researchers around things like disclosure review, and how not to do stupid stuff with the data. But we have largely left that, to date, to our PIs, which is

not working as well as one might hope' (SF007_1).

From the repository point of view, then, while there is no guarantee that researchers internalize a 'culture of confidentiality' by completing training modules, there is a comfort factor for repository staff when they do not have to trust solely on the word of the researcher that they have been exposed to the materials.

Reputation

The reputation category is separate from the identity and training categories because of how the components are viewed differently by repositories. This category is more important on its own merit for researchers who routinely and repeatedly access restricted data. As discussed in the next section, a researcher's reputation, evidenced through the interplay of qualifications, history of good data use, and status within the research community, contribute to the repository's development of trust. The researcher's reputation for good or bad data handling and overall 'trustworthiness' in the eyes of the repository staff were all highly influential as a basis for future access to data. As one interviewee described,

> 'I know, for example, when I worked in this RDC in Germany, there, whenever there was a breach, by default, all the other RDCs in Germany would get an email with the name and with what happened so every other agency in Germany is aware that there was a breach with one person' (SF005_1).

That the same components exist in the other categories – CV, institutional affiliation, completion of training modules, etc. – will be revisited during the discussion of the new trust model.

Project

The project category contains all other project related documentation reviewed by repository staff for scientific merit, data utility, and legal acquiescence that are not already included in the identity categories. This review process is focused on the project itself. While these components are usually reviewed by the same staff member, they are approved or denied based on project-specific considerations:

> 'we're only looking at, can you answer that? Or can you make any progress on that research question with our data? We don't assess, is that a good research question, you know, it's like, is the data suitable?' (SF006_1).

As with the other categories, the components are not universally required, though the three most common ones – project descriptions, signed data use agreements, and requests for specific datasets – are all used by more than half of the repositories. Other commonly requested items, with seven to eight repositories each, include data security plans, non-disclosure agreements, and funding source proof. The review for scientific merit is a review to ensure that the data requested will answer the researchers' questions, not a judgment on the research itself.

The New Model

Examining these categories through the lens of Boersma et al.'s (2003) three dimensions of trust, I found that while the project category does not have a distinct place in their model, the remaining three categories do. However, the categories do not perfectly fit into the dimensions, and therefore the basic model represented in Figure 1 does not accurately reflect how trust is developed in these repositories. The challenge lies in how competence-based and goodwill-based trust are defined in the original model, and which markers are used to demonstrate those dimensions in restricted data repositories. In this section, I present a revised model of trust development that resituates competence and goodwill within restricted data repositories.

Promissory-Based Trust

Boersma, Buckley, and Ghauri define the promissory-based trust dimension as the 'expectation that a party can be relied on to keep to agreements' (2003). In restricted data repositories, the markers of this dimension are the legal agreements (e.g., data use agreements, licenses, signed terms of use, etc.) which researchers, and their host institutions, sign. This dimension, as defined in the original model, is applicable to restricted data repositories. It applies to first-time restricted data users as well as returning users. In the interviews, several spoke of how repeat users with no history of bad data handling are assumed to be trustworthy to carry out the requirements of legal agreements. This demonstrates a close association between the promissory- and goodwill-based dimensions, where there is evidence of goodwill-based trust, promissory-based trust is assumed to a degree that it is not seen with users who do not have pre-existing good-will based trust.

Competence-Based Trust

Competence-based trust is defined as the expectation that the party 'will perform its role competently' (Boersma et al., 2003). For restricted data researchers, the role here reflects two different forms of competence, with different identifying markers, as discussed previously in the Training findings – research competence evidenced through professional and academic qualifications and institutional status, and restricted data competence evidenced through the successful completion of other training modules, courses, or certifications. Therefore, because of the insufficiency of the original dimension, I propose two new dimensions of competence.

Experience-based trust

The first new dimension I have called the 'experience-based trust.' This new dimension is the expectation that the researcher is experienced in conducting responsible and ethical research. The evidence for this includes the highest attained degree level, role at their institution, and record of prior-restricted data use. For example, an undergraduate student who has never used restricted data before would not be trusted with the same data set as a tenured faculty member with 15 years of experience without restrictions in mode of access and other protections that reflect their limited experience. The existence of the record of prior data use itself is indicative of previous experience, and thus an expected measure of experience working with

restricted data. But it is also indicative of trustworthiness to handle restricted data, as it would record both good data handling behaviour and bad data handling behaviour.

Competence-based trust

Unlike the original model, this 'competence-based trust' only refers to non-academic research competence as it is applicable to the use and protection of research data. I have re-defined this dimension as: the expectation that the party has internalised the culture of confidentiality required by restricted data providers, legislation, and data repositories. Evidence for this dimension are the successful completion of relevant restricted data-related trainings. This evidence of competence, while highly desired by most interviewees, would, I argue, follow the pattern of the experience-based trust. Researchers with a history of data use would be expected to also have a record of these training completions. The more exposure to concepts deemed important and valuable to the restricted data community, the more expectation there is that the researcher has internalized the important concepts.

Goodwill-Based Trust

Goodwill-based trust is used in the Boersma et al. model as the expectation that 'the party will behave honourably even where no explicit promises or performance guarantees have been made' (2003). For restricted data repositories, this means that, when the researcher is working unobserved with restricted data, they can be trusted to follow the rules, not attempt re-identification, and to properly supervise other data users. The development or lack of goodwill-based trust is in effect the reputation of the researcher within their research community, as understood by the repository. I recommend that the above definition is still valid in the case of restricted data repositories, and that the evidence for it be: the reputation in the researcher's community as shared between data repositories, good data stewardship, and a record of how well the researcher adhered to prior promissory documents. For first time data users, this dimension would carry less weight with repositories than it would with repeat data users who have built positive track records as trusted researchers.

The New Model

Therefore, for restricted data repositories, the development of trust can be modelled as the input of the Information, Training, and Reputation components into a review process, with the output a combination of the four dimensions of trust, the strength of which indicates a level of trust in the researcher. These dimensions are defined:

- **Promissory:** the expectation that the researcher can be relied on to keep to agreements;
- **Experience:** the expectation that the researcher is experienced in conducting responsible and ethical research;
- **Competence:** the expectation that the researcher has internalised the culture of confidentiality required by restricted data providers, legislation, and data repositories;
- **Goodwill:** the researcher will behave honourably even where no explicit promises or performance guarantees have been made.

This level of trust, when combined with the outcome of the project review, as depicted in Figure 3, reflects the increased level of scrutiny that goes into validating the identity of the person requesting access to data. It also reflects the dual review process used when validating a data access request: a determination of trust in the user, and a validation of the need for the data. The reflexive nature of the process is indicated by the input of 'successful completion of data use' into the review process for subsequent data access requests.

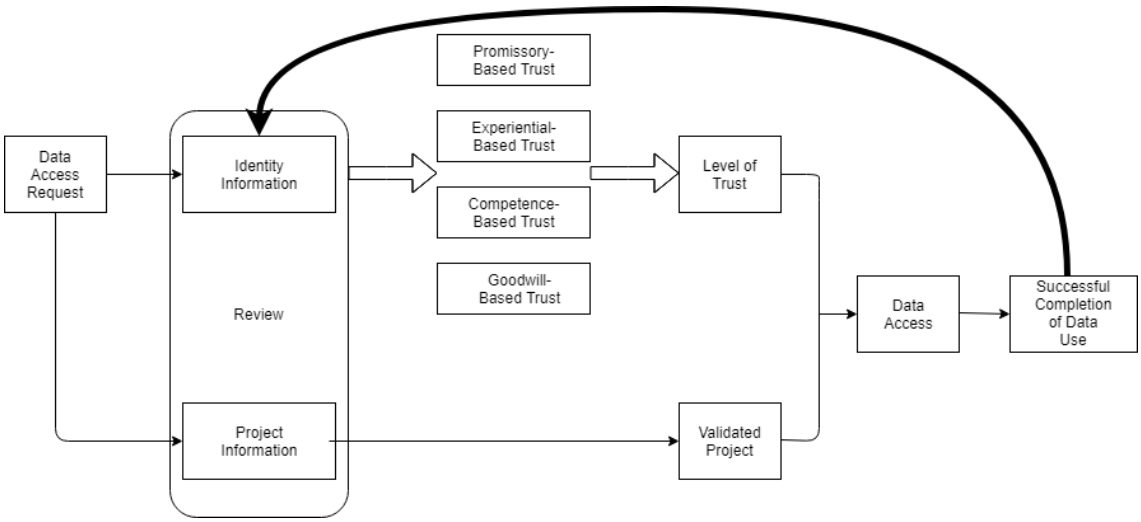

Figure 3. Model of the data access request process at restricted data repositories, depicting how trust is developed and combined with project validation to determine access to data.

This new model also demonstrates that the process of trust development in restricted data repositories is, unlike the serialized four stage process of IJV negotiations, occurs in two phases. When first-time data users submit an access request, the most important considerations in the review of the Identity, Training, and Reputation inputs is the ability to develop experience-based and competence-based trust (see Figure 4a). Promissory-based trust is more highly weighted in this phase because there is no indication of prior adherence to requirements. Therefore, the outcome of this review for first-time users is a level of trust based primarily on the relationship between promissory-, experience-, and competence-based trust, as there is not likely to be much pre-existing goodwill built up in this first phase.

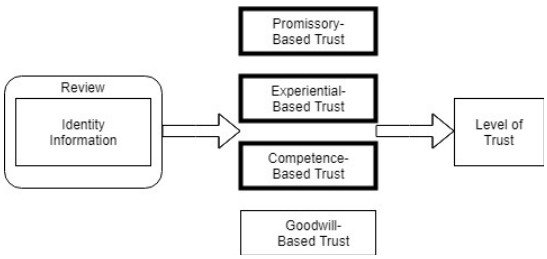

Figure 4a. The predominate dimensions of trust generated for first-time restricted data users.

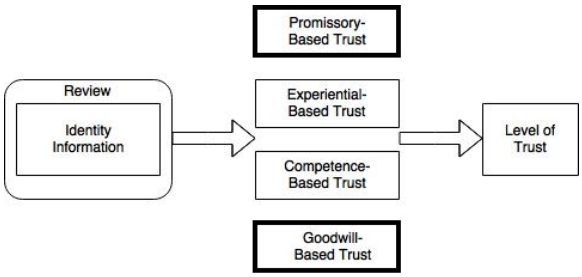

Figure 4b. The predominate dimensions of trust generated for repeat restricted data users.

The other phase (Figure 4b) involves researchers with prior restricted data experience. In this phase, while experience-based and competence-based indicators are considered, they are not as highly weighted in the evaluation process. With each successful completion of a restricted data project within the confines of the promissory documentation, a researcher's reputation within the restricted data community increases (or, with evidence of bad data handling, decreases). The stronger the reputation based on prior data use, the more trust is given to that researcher on their next data access request. The recursive nature of this process is what builds or diminishes a researcher reputation.

Discussion

To model the process of trust development at restricted data repositories, I have evaluated the fit of the Boersma et al. model of trust development to repository data access request processes. Due to conflated measures of competence, I presented a revised model of trust development that portrays the parallel processes of identity and project validation, and demonstrated, for restricted data repositories, the four dimensions of trust on which data users are evaluated. This model presents trust as the cumulative weight given to the four dimensions of trust, the strength of each dependent upon characteristics of the data reuser. This reflects the interaction-based nature of data repository-data user trust development over time. Prior work on trust development has focused on certifying repository and data trustworthiness (Donaldson and Conway, 2015; Yakel, Faniel, Kriesberg and Yoon, 2013; Yoon, 2014).

Current Developments

During and after this study, many data repositories investigated potential solutions to the challenge of increasing the utility of restricted data within legal and ethical boundaries (e.g., the 2018 implementation of the General Data Protection Regulation). Drawing on the Data Without Boundaries vision for a European Remote Access Network, the International Data Access Network (IDAN) is a collaboration between six European repositories in France, Germany, the Netherlands, and the UK to enable access to controlled access data remotely while in another country (Centre d'Accès Sécurisé aux Données, n.d.). The Australian Research Data Commons in 2018 merged together Australian resources to better support researchers and research data support (Australian Research Data Commons, 2019). While there will not likely be a one-size-fits-all solution to this challenge, different approaches and solutions are being developed.

ICPSR's solution is the Researcher Passport[1] [2], a digital access credential intended to be shareable between a consortium of restricted data repositories. This credential will contain within it verified information about researchers, minimizing the need for researchers to repeatedly provide the same information to data providers that must then be re-verified each time a researcher requests access to data. The Passport will incorporate an Open Badges protocol connecting the researcher to, among other identity components, the pre-data access training completion that this study identified as an important component of trust development. The first software development phase of the Passport was completed in 2018, and ICPSR MyData account holders are able to apply for the basic level Passport at this time. A Privacy Impact Assessment to identify potential privacy concerns was conducted in early 2019. The remainder of 2019 will be spent solidifying the Open Badges, the identification of and process for verification of identity information, and the alignment of the Passport components with data access restrictions.

Limitations

There are several limitations to the study that I carried out. First, I evaluated repositories around the world, and focused only on those that handle restricted data. Differences in legal requirements and international privacy regulations between, for example, the United States and Europe do not allow for the creation of a globally-accepted set of identity criteria, and so my recommendations must be viewed as generic recommendations to be fit within specific legal regimes. Second, the repositories I studied hold a wide variety of data types. While this breadth of repository types provides a range of identity criteria and validation processes to analyse, the resultant model requires further refinement to be applicable to individual repositories and their specific data security needs. Third, I only studied 23 data repositories and conducted nine interviews. Although many of these repositories represent the larger, more well known, and influential repositories, they are only a very small fraction of the digital research data repositories in the world. Only speaking to nine of those repositories means that I may have missed or left out other highly valued sections that may have changed the model or its indicators.

Conclusion

In this paper, I have utilized an economic model of transactional trust development to examine complex repository restricted data access requirements. In doing so, I have refined the original three-dimensional process model of trust development based on promissory, competence, and goodwill into a model that better reflects the unique identity requirements for researchers requiring restricted data. The new four-dimensional model situates the information about the individual from data access requests into the model as markers of the refined promissory, experience, competence, and goodwill-based trust. This allows for a greater understanding of the role of different types of experience that researchers are expected to have – experience based on

1 ICPSR Researcher Passport: https://www.icpsr.umich.edu/icpsrweb/content/about/researcher-credentialing.html

2 Development of the Researcher Passport is on-going under the National Science Foundation grant #1839868, "CICI: RDP: Open Badge Researcher Credentials for Secure Access to Restricted and Sensitive Data."

academic qualifications and training, and experience based on exposure to research community norms surrounding privacy, confidentiality, and restricted data management.

Acknowledgements

This study is a part of a larger Alfred P. Sloan Foundation-funded project (G-2016-7212) conducted through the Inter-university Consortium for Political and Social Research (ICPSR) to develop the policies and procedures that would allow for more efficient and effective access to restricted data for secondary analysis and research replication. I would like to thank Margaret Levenstein of ICPSR and Elizabeth Yakel of the University of Michigan School of Information for the opportunity to conduct this research, and for their valuable feedback. I would also like to thank the anonymous reviewers and the editor for their valuable feedback.

References

Akter, S., D'Ambra, J., & Ray, P. (2011). Trustworthiness in mHealth information services: An assessment of a hierarchical model with mediating and moderating effects using partial least squares (PLS). *Journal of the American Society for Information Science and Technology, 62*(1), 100–116. doi:10.1002/asi.21442

Australian Research Data Commons. (2019). ARDC Strategic Plan 2019-2023. Retrieved from https://ardc.edu.au/wp-content/uploads/2019/05/ARDC-Strategic-Plan-2019-2023.pdf

Boersma, M.F., Buckley, P.J., & Ghauri, P. N. (2003). Trust in international joint venture relationships. *Journal of Business Research, 56*(12), 1031–1042. doi:10.1016/S0148-2963(01)00315-0

Centre d'Accès Sécurisé aux Données. (n.d.). A collaborative project. Retrieved from IDAN International Data Access Network website: https://idan.network/

Corti, L. (2007). Re-using archived qualitative data – Where, how, why? *Archival Science, 7*(1), 37–54. doi:10.1007/s10502-006-9038-y

Dingwall, G. (2004). Trusting archivists: The role of archival ethics codes in establishing public faith. *The American Archivist, 67*(1), 11–30.

Donaldson, D.R. (2016). The digitized archival document trustworthiness scale. *International Journal of Digital Curation, 11*(1), 252–270. doi:10.2218/ijdc.v11i1.387

Donaldson, D.R., & Conway, P. (2015). User conceptions of trustworthiness for digital archival documents: User Conceptions of Trustworthiness for Digital Archival Documents. *Journal of the Association for Information Science and Technology, 66*(12), 2427–2444. doi:10.1002/asi.23330

Ertzscheid, O. (2016). What is digital identity? (H. Tomlinson, Trans.). Retrieved from http://books.openedition.org/oep/1388

Eschenfelder, K.R., & Johnson, A. (2014). Managing the data commons: Controlled sharing of scholarly data. *Journal of the Association for Information Science and Technology, 65*(9), 1757–1774. doi:10.1002/asi.23086

European Parliament. (2016). Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons With regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation).

Holdren, J.P. (2013). Increasing access to the results of federally funded scientific research [Memorandum]. Retrieved from Office of Science and Technology Policy website: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

Kelton, K., Fleischmann, K.R., & Wallace, W.A. (2008). Trust in digital information. *Journal of the American Society for Information Science and Technology, 59*(3), 363–374. doi:10.1002/asi.20722

Levenstein, M.C. (2019). The researcher passport: Improving data access and confidentiality protection, global, 2017-2018 — Interview protocol, qualitative data codeset, and exemplar transcript. Ann Arbor, MI: Inter-university Consortium for Political and Social Research. doi:10.3886/ICPSR37454.v1

Levenstein, M.C., Tyler, A.R.B., & Davidson Bleckman, J. (2018). The researcher passport: Improving data access and confidentiality protection: ICPSR's strategy for a community-normed system of digital identities of access (White Paper No. 1). Ann Arbor, M.I.: University of Michigan Inter-University Consortium for Political and Social Research.

Manhas, K.P., Page, S., Dodd, S.X., Letourneau, N., Ambrose, A., Cui, X., & Tough, S. C. (2015). Parent perspectives on privacy and governance for a pediatric repository of non-biological, research data. *Journal of Empirical Research on Human Research Ethics, 10*(1), 88–99. doi:10.1177/1556264614564970

National Institutes of Health. (2003). Final NIH statement on sharing research data (Notice No. NOT-OD-03-032). Retrieved from National Institutes of Health website: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html

Productivity Commission. (2017). Data availability and use (Inquiry Report No. 82). Canberra, Australia: Australian Government.

Rieh, S.Y. (2002). Judgment of information quality and cognitive authority in the web. *Journal of the American Society for Information Science and Technology, 53*(2), 145–161. doi:10.1002/asi.10017

Saldaña, J. (2016). The coding manual for qualitative researchers (3rd ed.). Beverly Hills, C.A.: Sage Publications.

Schensul, J.J., & LeCompte, M.D. (2013). Essential ethnographic methods: A mixed methods approach. Lanham, M.D.: AltaMira Press.

SpringerNature. (2017). Availability of data, material and methods. Retrieved from Nature.com website: http://www.nature.com/authors/policies/availability.html

Strauss, A., & Corbin, J. (1990). Basics of grounded theory methods. Beverly Hills, C.A.: Sage.

Sullivan, C. (2012). Digital identity and mistake. *International Journal of Law and Information Technology, 20*(3), 223–241. doi:10.1093/ijlit/eas015

Tenopir, C., Dalton, E.D., Allard, S., Frame, M., Pjesivac, I., Birch, B., … Dorsett, K. (2015). Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PLOS ONE, 10*(8), e0134826. doi:10.1371/journal.pone.0134826

Wynne, B. (1992). Misunderstood misunderstanding: Social identities and public uptake of science. *Public Understanding of Science, 1*, 281–304.

Yakel, E., Faniel, I., Kriesberg, A., & Yoon, A. (2013). Trust in digital repositories. *International Journal of Digital Curation, 8*(1), 143–156. doi:10.2218/ijdc.v8i1.251

Yoon, A. (2014). End users' trust in data repositories: Definition and influences on trust development. *Archival Science, 14*(1), 17–34. doi:10.1007/s10502-013-9207-8