

Restoring Trust Relationships within the Framework of Collaborative Digital Preservation Federations

**Robert H. McDonald (Indiana University) and
Tyler O. Walters (Georgia Institute of Technology)**

Indiana University, Herman B. Wells Library 234
1320 East 10th Street, Bloomington, IN 47405-3907
Georgia Institute of Technology, Library and Information Center
704 Cherry Street, Atlanta, GA 30332-0900
{robert} at indiana.edu, {tyler} at gatech.edu

Abstract

The authors extend their process for creating and establishing trust relationships to include steps for restoring trust relationships after catastrophic events. Part of this model will include best practices for business continuity relationships and will integrate trust models from Holland and Lockett and Ring and Van de Ven. These models provide key frameworks for understanding how trust can be utilized as collaborative starting points as well as recovery points from physical natural disaster or critical data loss. Additionally, the authors will present findings from an audience poll conducted on May 15, 2009 during the presentation of this topic at the 4th *International Conference on Open Repositories* with additional conclusions and commentary.

Introduction

As more research, educational, and cultural institutions come to realize the enormity and complexity of work required to store, preserve, and curate large amounts of their unique digital information, many will turn to establishing cooperative partnerships for leveraging existing mass-storage capacity or utilizing 3rd party data curation service providers to help satisfy their needs for a redundant and secure digital preservation system (Jordan, McDonald, et. al. 2008). The concept of trust and its manifestation between institutions as an essential element in designing digital preservation systems – both technical and organizational – is critical and appears in the organizational level needs of the CRL/NARA-RLG Trustworthy Repositories Audit and Certification (TRAC): Criteria and Checklist. Trust can be defined simply as “relying upon or placing confidence in someone or something...” (www.dictionary.com). With regard to preservation in digital libraries and archives, trust means that we rely upon the organizations or institutions maintaining the digital library or archives to sustain the

information deposited in it, and that this information remains authentic, reliable, and unchanged over time and across technologies. We trust that the institutional actions taken upon the digital library and the content held can be trusted to serve these goals. To achieve this, as we look at partner institutions who are participating in preserving our own institution’s digital content, we are seeking to answer whether or not their actions with our material are trustworthy. Trust is always an underlying, critical factor impacting the success or failure of inter-institutional relationships. The concept of trust is imbued in everything we do as digital library and archives professionals, especially in an inter-institutional, cooperative setting.

Increasingly, federations of institutions and organizations are being formed to devise strategies and systems to preserve digital information. The choice of the word “federations” is significant because it aptly describes what these institutions are doing. “Federation” can be defined as “people, societies, unions, states etc. joined together for a common purpose.” “...a federated body formed by a number of nations, states, societies, unions, etc., each retaining control of its own internal affairs.” (www.dictionary.com). According to these definitions, a federation is unique in that the individual institutions comprising it continue to “retain control of its own internal affairs,” while at the same time they are coming together to solve a common need. The phrase “distributed digital preservation federations” is being used increasingly to describe cooperatives of geographically-dispersed institutions who are banding together to form solutions to the digital preservation problem. Identifying and analyzing successful federation models as well as human practices that foster inter-institutional trust development are salient to the work of building distributed digital preservation federations.

Frameworks for Trust

Within any model for distributed preservation, people, organizations, and the inter-institutional federations between them must have a formal mandate for “trust.” This type of formalized trust has been previously identified within the preservation literature from both a contractual (Berman et. al., 2008); evidence based methodology (Ross and McHugh, 2006), and organizational structure analysis (McDonald and Walters, 2007). In order for this trust model to succeed when applied to coordinated or federated digital preservation organizations, each autonomous entity must receive adequate preservation services while retaining appropriate autonomy for its primary institutional organization. The authors will delve further into examining what institutional and personal characteristics, principles, and building blocks must be present to foster and sustain trust in an inter-organizational model such as digital preservation federations. They will describe and discuss the dynamics of such a model in light of a disaster or critical data loss.

Holland and Lockett. In the first model examined here for transaction-based trust relationships we have identified one set forth by Holland and Lockett which looks at virtual organizational models. The prime motivator in this model is the idea of business and commerce being motivated by many complex partnerships in the supply chain in order to conduct business at a global scale. Much like the types of international trust relationships that digital preservation cooperatives seek, this virtual environment is built upon indicators of trust. In cases of digital curation federation this trust process must be formalized in some form of agreement whether it is one that is legally binding or intended for production support or for organizational clarity.

Holland and Lockett devise five hypotheses which will be telling in the long-run as to how effective virtual organizations can be in managing national and international preservation efforts. These hypotheses are as follows (Holland and Lockett, 1998):

Hypothesis 1: Virtual organizations will develop quicker and easier where the level of subjective trust between the different economic partners is high.

Hypothesis 2: The importance of subjective trust in determining the success of virtual organizations is contingent on the risk of failure and the importance of the outcome.

Hypothesis 3. Shared information systems amongst economic partners involved in some form of virtual organization will serve to speed up the trust/distrust development process.

Hypothesis 4. International differences in dispositional

trust will become less important than situational context in determining the level of subjective trust as shared information systems enable the free flow of performance information between separately owned economic partners.

Hypothesis 5:

In business markets, virtual organizations will be characterized by long-term relationships and stability rather than transient relationships to support unique projects or electronic markets.

Ring and Van de Ven. This (1994) model is designed to examine cooperative inter-organizational relationships (IORs) and the frameworks they utilize in formal, legal, and informal social-psychological processes when negotiating and executing their business activities. Ring and Van de Ven further focus upon and explore how and why cooperative IORs emerge, evolve, and dissolve. They assert that their findings enlighten our understanding of the transactional cost economics of business being conducted through cooperative IORs as well as other aspects of business relationship development. Their modeling can help in understanding the characteristics of digital preservation federations’ lifecycle stages as these efforts are initiated, ascend, and mature. Since many federated curation collaboratives operate in a non-profit status as an IOR, it is important that the governance models put in place to support these IOR’s provide for clarity in operational, procedural, and policy initiatives. The Ring and Van de Ven model provides a framework with which the IOR can be instantiated.

Proposition 1: Congruent sense making among parties increases the likelihood of concluding formal negotiations to a cooperative IOR.

Proposition 2: Congruent psychological contracts among parties increases the likelihood of establishing formal commitments to a cooperative IOR.

Proposition 3: If the individuals assigned to a cooperative IOR do not change, personal relationships increasingly supplement role relationships as a cooperative IOR develops over time.

Proposition 4: Informal psychological contracts increasingly compensate or substitute for formal contractual safeguards as reliance on trust among parties increases over time.

Proposition 5: When the temporal duration of inter-organizational relationships is expected to exceed the tenure of agents, informal understandings and commitments will be formalized.

Proposition 6: As the temporal duration of a cooperative IOR increases, the likelihood decreases that parties will

terminate the relationship when a breach of commitments occurs.

Proposition 7: When significant imbalances between formal and informal processes arise in repetitive sequences of negotiation, commitment, and execution stages over time, the likelihood of dissolving the cooperative IOR increases.

Case Study for Disaster

For illustrative purposes, the authors will apply the trust frameworks and recovery processes we just outlined to a fictional shared data repository that serves the bandwidth and geographical location of the Southeastern United States. This fictional institution, the SRAST, or Southeast Regional Advanced Storage Trust, provides digital repository access and preservation services and mass-scale storage for research libraries and data centers in the U.S. southeast. The SRAST is based in the research triangle of North Carolina. The data that SRAST preserves is replicated and backed up across the national lambda rail at a similar data center based in north Georgia. The Georgia center's purpose is as a dark storage and preservation node only; all access to researchers is provided by SRAST in North Carolina.

The Fictional Case: The SRAST digital repository service was severely tested when a Category 5 hurricane struck directly at the coast of North Carolina wiping out power for most of the state. The computing facility was significantly compromised with major structural damage sustained. The building will need a three-month recovery time with only some access in the first couple of weeks. Meanwhile, 400 miles away in Georgia, the hurricane brought 15 inches of rain within a three-hour period. The north Georgia data-center facility was submerged in water from a local river with acres of mud flooding the city and computing facility. Power was lost for 48 hours.

Research data stored on servers via the digital repository service in North Carolina and Georgia were lost. Much of the data also was backed up on tape in Georgia (a rotational tape backup which leaves at least 48 hours worth of data loss). Also, the tape storage room in Atlanta was hit by the water and mud that entered the building. Many tapes were damaged by the water and mud, while others were subjected to high humidity levels, all compromising the tapes' structure. Only a portion of the tapes could be recovered, with only portions of research data retrieved.

While this is fictional, the 1989 Category 5 hurricane (Hugo) on the U.S. Mid-Atlantic coast and the 2005 hurricanes on the U.S. Gulf-Coast have shown us that this type of scenario is within the realm of possibility. In this case study, research data kept via a digital preservation service definitely has been lost. What is this collaborative

storage facility to do to regain trust from its customer institutions? The authors will consider the three recommended trust recovery steps we articulate.

Applying Trust Frameworks After Disaster Strikes

Of course, having a framework for inter-organizational trust prior to disaster or loss of continuity is critical and engaging the smaller groups that direct the work of the IOR or VO (virtual organization) is important. However, there is always the chance that a critical loss or disaster will occur prior to institutional membership or during a natural disaster. In order to re-establish trust after such a situation occurs the authors feel that a multi-tiered approach must be used in order for a full recovery to occur. This approach involves three components including a 1.) Initial Public Communiqué 2.) Transparent Accounting of Circumstances that Led to Data Loss and Recovery. 3.) Independent Audit.

Initial Public Communiqué. After the physical facilities have been stabilized and the loss of data has been understood and verified, SRAST should initiate a public communiqué that also goes directly to its customer institutions. The communiqué must publicly acknowledge the disaster that occurred and the extent of damage to facilities and IT systems, as well as the extent to which research data was lost. The means to communicate with SRAST personnel and management with responsibility for specific areas (i.e. facilities, IT systems, data management, and customer relations) should be given. The intent of the communiqué is for SRAST's management to demonstrate openness, transparency, and authenticity in SRAST's communications with the public, its partners, and customers. Holland and Lockett address the importance of building trust through communicating information. Their Hypothesis 3 holds that sharing information systems for the purpose of communication between partners will strengthen and advance trust in the virtual organization. Therefore, our recommendation of an initial public communiqué is consistent with this hypothesis, as a first step toward keeping trustworthy business relationships together. It is extremely important to have a thorough communication plan in place that is expressed in an MOU and in an annual renewal statement. While strategic partners need to know what has happened during a crisis, they may not want this information to be public knowledge. However, disclosing the occurrence establishes public credibility in the face of disaster. In the business continuity literature, there are several frameworks that could be utilized for this planning scenario but all have multiple facets for near-term emergency communication and long-term multi-tiered communications responses (Childs, 2008).

Transparent Accounting of Circumstances that Led to Data Loss and Recovery. An independent third party should be employed to investigate the conditions that created the vulnerabilities that SRAST found itself with during the natural disaster. Much like the general audit described in Step 3 in the trust recovery process, this transparent accounting focuses expressly upon the strengths, weaknesses, and opportunities in SRAST's IT, organizational, and accountability systems, and documents them. The final report of findings is a public document that is transmitted to SRAST's management as well as the customer/partner institutions. It becomes the basis for a full independent audit that begins the process of improving and reengineering SRAST's processes, policies, and infrastructures to rebuild its services and the trust of its customers/partners. This shared system of transparent accountability highlights Hypotheses 3 and 4 in the Holland and Lockett Model (1998). In Hypothesis 4, creating and sharing performance information between the parties is most important when a situation occurs such as a disaster and consequent data loss. In our recommendation, we suggest the independent third-party be used because it is in the best position to provide objective, accurate, and unbiased information about the occurrence and about SRAST's actions to disclose, inform, and recover from it. Also, if trust was in place to begin with, then SRAST should move to retain that trust by supplying its clients with disaster-related performance information as expeditiously as possible. This should be done of SRAST's own initiative, without it being requested by the clients. Such actions are consistent with Ring and Van de Ven's Proposition 4, which suggests that informal agreements and actions will increasingly be relied upon as the reliance on informal trust among the parties' increases over time.

Independent Audit. An independent third party with appropriate expertise needs to audit SRAST's overall digital repository practices, using techniques such as TRAC and DRAMBORA. A full report needs to be made on the veracity of SRAST's technical approaches including their apparent strengths, weaknesses, and opportunities, with a focus on its policies and procedures, organizational methods to promote transparency, documentation practices, and formalized inter-organizational relationships. This report should be released to SRAST and the customer/partner institutions involved in the trust. SRAST should use the findings of this report as a pathway to rebuild its IT systems, management and accountability systems, inter-organizational relationships, and services. This type of independent accounting over time will lengthen the long-term involvement of partners and will strengthen relationships that otherwise could be jeopardized.

In their Proposition 7, Ring and Van de Ven describe how the risk of rising distrust and disillusionment within the

inter-organizational cooperative comes from an imbalance between formal and informal relationship maintenance processes. In Step 2, we recommended an independent third party to conduct a transparent accounting of the circumstances that led to the data loss. This step contains a balance of formal and informal processes; the formal being the report generated by the independent entity and the informal being SRAST's self-initiated actions to share performance information. In Step 3, we are mindful of this need to balance the formal and informal processes to maintain and even restore trust when a data loss occurrence has damaged it. The independent audit utilizing approaches such as TRAC and DRAMBORA, and the resulting report and recommendations, represent a formal trust-building process in the relationship between SRAST and its customers/partners. Informal processes are represented by SRAST's actions related to open and transparent communication after the report. Self-initiating to involve the customers in SRAST's post-report action steps, maintaining meetings and talking with the customers' representatives are examples of informal communicative processes. The formal, independent auditing process and the informal communication and "buy-in" processes help balance the formal and informal to restore trust, as asserted by Ring and Van de Ven.

Presentation and Audience Polls from OR09

During the authors' recent presentation at the 4th International Conference on Open Repositories on this topic, an audience response system (ARS) was used to gauge both experience for the disaster scenario and the relevance of the material both for the speakers and for the audience. The authors utilized this system in order to obtain active comprehension from the audience (Caine and Robinson, 2008); however, in addition to the active learning comprehension the ARS provided interesting anonymous response information from a diverse group of repository professionals. This information would be an interesting basis for a larger poll of repositories to look at best practices for disaster recovery and more importantly for re-establishing trust after such a catastrophic event.

The following questions were posed using the free Web based ARS textthemob.com. The audience could respond via SMS text message or Web based browser and had the options for *Yes*, *No*, or *Do Not Know*.

The audience was comprised of at least 200 people and while not everyone chose to respond to the poll, the outcomes do provide some areas of exploration for future discussion in the repository community.

The following questions were posed to the audience with corresponding results.

Question 1 - Do you host your repository and its storage in your own data center?

Responses – Yes-24 No-2 Do Not Know-0

Question 2 - Do you host your repository and its storage in a co-located data center facility that is managed for you?

Responses – Yes-12 No-15 Do Not Know-1

Question 3 - Do you have geographically dispersed replicated storage space for your institutional repository?

Responses – Yes-8 No-21 Do Not Know-1

Question 4 - Do you have a federated partnership with other institutional repositories for data replication or preservation?

Responses – Yes-9 No-23 Do Not Know-0

Question 5 - Have you ever had hardware failures within your institutional repository infrastructure?

Responses – Yes-17 No-10 Do Not Know-0

Question 6 - Have you ever lost data from your institutional repository?

Responses – Yes-3 No-21 Do Not Know-4

Audience Conclusions

From our audience response questions we believe that Question 4 shows an increase in federated partnerships; however, without more formal polling this assumption is only based on a working knowledge of these partnerships over the last decade. While it may not be clear that there is an emerging and significant interest in creating collaborative digital preservation services, the authors believe that this type of federation will become more predominant within the non-profit repository community. Question 4 does show that 28% of respondents are involved in a federated partnership. While this is not a majority, we interpret this as a rise in using multi-institutional, federated approaches. Also, in Question 2 responses, over 42% replied they are involved in a co-located facility that is managed by another organization for them. Institutions involved in these collaborative services also need to investigate further how they could host their content in geographically dispersed locations. Or have redundancy that can mitigate disaster scenarios that are described within this article. Documenting and instantiating procedures, workflows and policies for what would happen during any catastrophic systems failure as well as routine day to day hardware failures or human

management errors are just as necessary when a repository is in a co-located facility as when the system is managed by an IOR. The trust rubrics referenced in this paper will help with that documentation and with any other memorandum that would serve to document a federated digital preservation service that leverages data handling capabilities at multiple institutions.

Overall Conclusion

Data loss in complex systems, whether through natural disaster or more likely through human error, is inevitable. Recovering from these phenomena is an organizational challenge that will become an ever-increasing dilemma for research, educational, and cultural organizations as their artifacts become born-digital in nature. Models such as the ones put forth by Ring and Van de Ven as well as Holland and Lockett are designed to examine cooperative inter-organizational relationships (IORs) and the frameworks they utilize in formal, legal, and informal social processes when negotiating and executing trusted business activities. Applying these trust frameworks in initial agreements or service contracts between service collaborators will enhance trust relationships in cases both for standard operating procedures as well as in the case of disaster and data loss. Understanding the complex human and organizational constructs for trust relationships will aid in developing a recovery construct that responds to more than just the technical constraints of risk within a collaborative curation environment,

Taking steps to demonstrate that a repository service like the fictional SRAST is transparent, and therefore credible, is crucial to disaster recovery and for the potential of re-instituting trust between organizations. Its organizational management team must demonstrate that it comprehends the disaster, is accurate and reliable with the information it shares widely, and allows independent third parties to verify the occurrence, analyze repository operations, and recommend action steps to reestablish trust in the repository service. The three steps described by the authors address these goals and should result in a digital repository service that can overcome lapses in trust relationships in order to provide sustainable repository services.

As noted from our audience response system questions, federated repository services are taking on important roles for groups involved in collaborative digital preservation and with that in mind, repositories will need to seek geographically dispersed redundant data systems for their repository content and will want to consider the trust relationships involved during both catastrophic systems failures as well as those that involve typical human error and hardware systems failures.

Authors

Robert H. McDonald is the Associate Dean for Library Technologies, Associate Director for the Data to Insight Center in the Pervasive Technology Institute, and Executive Director of Kuali OLE at Indiana University in Bloomington, IN.

Tyler O. Walters is the Associate Director for Technology and Resource Services at the Georgia Institute of Technology Library and Information Center in Atlanta, GA.

References

Berman, F., A. Kozbial, R.H. McDonald, B.E.C. Schottlaender. 2008. The Need to Formalize Trust Relationships in Digital Repositories. *Educause Review* 43(3).

<<http://connect.educause.edu/library/erm0835>>.

Childs, Melody. 2008. Hazards and Hurricanes: Hallmarks of IT Readiness, Response, and Recovery, *ECAR Research Bulletin*. v. 2008:21.

<<http://connect.educause.edu/Library/Abstract/HazardsandHurricanesHallm/47437>>.

Dictionary.com. 2008. Definition of *trust*. Accessed on 14th January, 2009. <<http://dictionary.reference.com/browse/trust>>.

Dictionary.com. 2008. Definition of *federation*. Accessed on 14th January, 2009. <<http://dictionary.reference.com/browse/trust>>.

Holland, C.P. and A.G. Lockett. 1998. Business Trust and the Formation of Virtual Organizations. *Proceedings of the Thirty-First Hawaii International Conference on System Sciences*, v. 6: 602-10.

Jordan, C., R.H. McDonald, D. Minor, and A. Kozbial. 2008. Cyberinfrastructure Collaboration for Distributed Digital Preservation. *IEEE Fourth International Conference on eScience*. Indianapolis, IN.

McDonald, R.H. and T. O. Walters. 2007. Sustainability Models for Digital Preservation Federations. *Proceedings of DigCCurr 2007: An International Symposium in Digital Curation*. <<http://hdl.handle.net/1853/14442>>.

Ring, P.S. and A. Van de Ven. 1994. Development Processes of Cooperative Interorganizational Relationships. *Academy of Management Review*, Vol. 19(1): 90-118.

RLG-NARA Digital Repository Certification Task Force. 2007. *Trustworth Repositories Audit and Certification: Criteria and Checklist*. <<http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=91>>.

Ross, S., and A. McHugh. 2006. The Role of Evidence in Establishing Trust in Repositories. *D-Lib Magazine* 12(7/8). <<http://www.dlib.org/dlib/july06/ross/07ross.html>>.

Cain, C. and J. Robinson. 2008. A Primer on Audience Response Systems: Current Applications and Future Considerations. *American Journal of Pharmaceutical Education*, Vol. 72(4): 77. Pubmed Central <PMCID: PMC2576416>.