The Social Construction of Risk in Trustworthy Digital Repository Certification

Prof. Rebecca D. Frank, PhD
Assistant Professor
Einstein Centre Digital Future
Berlin School of Library & Information Science
rebecca.frank@hu-berlin.de | @frankrd
I conduct research in the areas of digital preservation, digital curation, and data reuse (including open data), focusing on social and ethical barriers that limit or prevent the preservation, sharing, and reuse of digital information.
Background
## Trustworthy Digital Repository (TDR) Certifications

<table>
<thead>
<tr>
<th>Certification</th>
<th>Locations of Certified Repositories</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoreTrustSeal (Data Seal of Approval, World Data Systems)</td>
<td>Africa, Asia, Australia, Europe, North America</td>
</tr>
<tr>
<td>Trustworthy Repositories Audit &amp; Certification (TRAC) (ISO 16363)</td>
<td>Asia, North America</td>
</tr>
<tr>
<td>nestor Seal for Trustworthy Digital Archives (DIN 31644)</td>
<td>Europe</td>
</tr>
</tbody>
</table>
Study 1: TRAC (ISO 16363)
“A trustworthy digital repository will understand threats to and risks within its systems. Constant monitoring, planning, and maintenance, as well as conscious actions and strategy implementation will be required of repositories to carry out their mission of digital preservation.”

(ISO 16363, 2012)
Classical Approach to Risk

(probability of event) * (magnitude of consequences) = risk
“The classical risk approach assumes that it is possible to define and assess risks. The assumption that risks can be objectified and calculated has met with a lot of criticism.”

(van Est et al., 2012, p. 1075)
A Model for the Social Construction of Risk

- Communication
- Vulnerability
- Complexity
- Expertise
- Uncertainty
- Trust
- Organizations

Risk
Perceptions of risk vary depending on the way in which information about those risks is communicated, including the source, method, channel, and means of communication. (e.g. Bostrom, 2014; Kasper & Kasper, 1996)
A Model for the Social Construction of Risk

Communication
e.g. Bostrom, 2014; Kasper & Kasper, 1996

Complexity
e.g. Perrow, 1999; Wilkinson, 2001

Expertise
e.g. E. Vaughan & Seifert, 1992; Wynne, 1992

Organizations
e.g. Hutter, 2005; D. Vaughan, 1996

Trust
e.g. Nelkin, 1989; Wildavsky & Dake, 1990

Uncertainty
e.g. Starr, 2003; van Est et al., 2012

Vulnerability
e.g. Murphy, 2006; Olofsson et al., 2014
1. How do auditors and repository managers conceptualize risk in the context of a TRAC audit?

2. What are the differences and similarities by which auditors and repository managers understand risk as it has been communicated by the developers of the ISO 16363 standard?

3. To what degree do the following seven factors which influence risk perception come into play in the audit process: communication, complexity, expertise, organizations, trust, uncertainty, and vulnerability?
Semi-Structured Interviews
- Standard Developers (n=11)
- Auditors (n=10)
- Repository Staff (n=21)

Document Analysis
- ISO 16363 Standard (including audit checklist)
- Audit Reports
- Repository documentation
Findings
“Do I think that large amounts of people really understand how risk is constructed and what it means? No. … I think it’s relatively easy to get information about solutions and how things are implemented, and it’s harder to put that in a framework where you’re measuring the likelihood if it happening against the potential of it happening, and what the downsides are there, and how you tie specific numbers to that.”

(Repository Staff 18)
Types of Risk

- Financial
- Legal
- Organizational Governance
- Repository Processes
- Technical Infrastructure
Types of Risk

Financial

Legal

Organizational Governance

Repository Processes

Technical Infrastructure

Created by jai from Noun Project
Created by Creative Stall from Noun Project
Created by Evgeniy from Noun Project
Created by Creative Stall from Noun Project
Created by Denis Shapovalov from Noun Project
Constructing Risk Through a TRAC Audit
3.1.2.1 The repository shall have an appropriate succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.

Supporting Text
This is necessary in order to preserve the information content entrusted to the repository by handing it on to another custodian in the case that the repository ceases to operate.

Examples of Ways the Repository Can Demonstrate It Is Meeting This Requirement
Written and credible succession and contingency plan(s)
TRAC Audit
Performing Trustworthiness

“What is really going to be the reason repositories are at risk, is almost all around **having enough money** to take care of the material . . . a succession plan to move it someplace else, where the community isn’t going to have enough money to take care of it. Or there’s going to be a, someone who magically dumps money on the secondary repository. **Why couldn’t they dump money on the first repository?** I mean, it’s just, I don’t know. **It doesn't make sense.**”

(Repository Staff 12)
Different Perspectives

- Auditors
- Standard Developers
- Repository Staff

Risk
- Communication
- Vulnerability
- Complexity
- Uncertainty
- Expertise
- Trust
- Organizations
Implications: Risk in Digital Preservation

Technical
Economic
Organizational
Implications: Risk in Digital Preservation

Technical
Economic
Organizational
+ Social
A Revised Model for the Social Construction of Risk
Study 2: CoreTrustSeal (in progress)
<table>
<thead>
<tr>
<th>Certification</th>
<th>Locations of Certified Repositories</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoreTrustSeal (Data Seal of Approval, World Data Systems)</td>
<td>Africa, Asia, Australia, Europe, North America</td>
</tr>
<tr>
<td>Trustworthy Repositories Audit &amp; Certification (TRAC) (ISO 16363)</td>
<td>Asia, North America</td>
</tr>
<tr>
<td>nestor Seal for Trustworthy Digital Archives (DIN 31644)</td>
<td>Europe</td>
</tr>
</tbody>
</table>

**Diagram:**
- Auditors
- Standard Developers
- Repository Staff
The goal of this research is to understand how stakeholders in the CoreTrustSeal trustworthy digital repository certification process construct their understanding of risk.
Research Questions

● How do stakeholders in the CoreTrustSeal audit/certification process construct their understanding the concept of risk?
● What are the similarities and/or differences in how these stakeholder groups understand risk?
● How do the following factors affect an audit: communication, complexity, expertise, organizations, political culture, trust, uncertainty, vulnerability?
  ○ Which factors emerge at the individual level, and which at the social/group level?
● How do stakeholders perceive the value of the CTS certification process?
Mixed Research Methods

Phase 1: Survey

Phase 2: Interviews & Document Analysis

Phase 3: Observations

Phase 4: Workshop
Survey Distribution

● Sent via email to one representative from each CTS certified repository by the CTS Board
  ○ (or legacy certifications: DSA, WDS)

● Response rate: 54%
  ○ 163 certified repositories at the time of the survey
  ○ 88 complete responses
Survey: Preliminary Findings
Which certification(s) has your repository achieved? Select all that apply. (n=88)

<table>
<thead>
<tr>
<th>Certification</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoreTrustSeal</td>
<td>61</td>
</tr>
<tr>
<td>Data Seal of Approval</td>
<td>27</td>
</tr>
<tr>
<td>World Data Systems</td>
<td>44</td>
</tr>
<tr>
<td>Other</td>
<td>8</td>
</tr>
</tbody>
</table>

- CTS only: 20
- CTS + DSA: 20
- CTS + WDS: 25
- CTS + DSA + WDS = 4
### Survey Participants: Role

<table>
<thead>
<tr>
<th>Role</th>
<th>What is your <em>current</em> role?</th>
<th>What was your role at the time of the audit?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration</td>
<td>28</td>
<td>26</td>
</tr>
<tr>
<td>Digital Preservation</td>
<td>23</td>
<td>25</td>
</tr>
<tr>
<td>IT</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Other</td>
<td>31</td>
<td>27</td>
</tr>
<tr>
<td>No answer</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
### Survey Participants: Audit Process

**What was your role in the certification process? Select all that apply. (n=88)**

<table>
<thead>
<tr>
<th>Role</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prepared Documentation</td>
<td>80</td>
</tr>
<tr>
<td>Interacted directly with reviewers or auditors</td>
<td>54</td>
</tr>
<tr>
<td>Other</td>
<td>16</td>
</tr>
</tbody>
</table>
### Repository Staff & Audit Reviewers

**Have you served as a reviewer for another repository’s CTS, DSA, or WDS certification?**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>27</td>
</tr>
<tr>
<td>no</td>
<td>57</td>
</tr>
<tr>
<td>other</td>
<td>4</td>
</tr>
</tbody>
</table>

**If yes, how many audits have you participated in as a reviewer?**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>7</td>
</tr>
<tr>
<td>3-5</td>
<td>9</td>
</tr>
<tr>
<td>6+</td>
<td>12</td>
</tr>
<tr>
<td>total</td>
<td>28</td>
</tr>
</tbody>
</table>
Reminder: Types of Risk from TRAC Study

- Financial
- Legal
- Organizational Governance
- Repository Processes
- Technical Infrastructure
# Types of Risk

The items below indicate potential sources of risk for digital repositories, that could interfere with a repository’s ability to preserve digital content long-term. Please rank them in order from most to least significant.

<table>
<thead>
<tr>
<th></th>
<th>Most Significant</th>
<th>Significant</th>
<th>Neither Significant nor Insignificant</th>
<th>Slightly Insignificant</th>
<th>Least Significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finance (e.g. funding sources)</td>
<td>56</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Legal (e.g. rights)</td>
<td>12</td>
<td>2</td>
<td>16</td>
<td>13</td>
<td>45</td>
</tr>
<tr>
<td>Organizational Infrastructure (e.g., staffing)</td>
<td>15</td>
<td>33</td>
<td>25</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Repository Processes (e.g., digital object management)</td>
<td>4</td>
<td>11</td>
<td>17</td>
<td>35</td>
<td>21</td>
</tr>
<tr>
<td>Technical Infrastructure</td>
<td>11</td>
<td>22</td>
<td>22</td>
<td>21</td>
<td>12</td>
</tr>
</tbody>
</table>
Listed below are the three main sub-sections of the CoreTrustSeal requirements for certification. Please rank them in order of importance for determining whether a repository is able to preserve digital content long-term.

<table>
<thead>
<tr>
<th></th>
<th>Most Important</th>
<th>Important</th>
<th>Least Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizational</td>
<td>46</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Infrastructure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital Object</td>
<td>31</td>
<td>35</td>
<td>22</td>
</tr>
<tr>
<td>Management</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td>11</td>
<td>32</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In the Organizational Infrastructure section of the CTS requirements, Requirement R3 states:

“The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.”

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>In your opinion, is a repository having a continuity plan necessary for long-term preservation of digital content?</td>
<td>40</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Do you think that the items described in the guidance section above provide a sufficient evidential basis to demonstrate that a repository has met Requirement R3?</td>
<td>43</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Do you think that meeting Requirement R3 as described above would demonstrate a repository’s trustworthiness with regard to long-term preservation of digital content?</td>
<td>32</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>
Conclusions
(risk identification) + (documented mitigation strategy) = TDR certification

(risk identification) + (documented mitigation strategy) + (________) = trustworthy
Next Steps & Future Work

- **Survey**
  - Qualitative analysis of write-in responses

- **Interviews**
  - Continue interviews with stakeholders in the CTS certification process

- **Future research:**
  - Further refine the Model for the Social Construction of Risk in Digital Preservation
  - Investigate how risk is constructed among repositories that conduct self-audits (systems with standard developers and repository staff, but no formal auditors)
https://doi.org/10.1002/asi.24247 Available Open Access: http://hdl.handle.net/2027.42/149147

Acknowledgments

● Funding:
  ○ TRAC: University of Michigan Rackham Graduate School
  ○ CTS: Einstein Centre Digital Future

● Research Assistants
  ○ Megh Marathe & Carl Haynes (University of Michigan)
  ○ Laura Rothfritz & Maricia Mende (Humboldt-Universität zu Berlin)

● Participants
  ○ The Center for Research Libraries
  ○ Primary Trustworthy Digital Repository Authorisation Body Ltd.
  ○ Canadiana.org, Chronopolis, CLOCKSS, HathiTrust, Portico, & ScholarsPortal
  ○ CoreTrustSeal Board
The Social Construction of Risk in Trustworthy Digital Repository Certification

Prof. Rebecca D. Frank, PhD
Assistant Professor
Einstein Centre Digital Future
Berlin School of Library & Information Science
rebecca.frank@hu-berlin.de | @frankrd
References

References