Repository Staff Perspectives on the Benefits of Trustworthy Digital Repository Certification

Rebecca D. Frank

University of Tennessee, Knoxville, rfrank7@utk.edu

Follow this and additional works at: https://trace.tennessee.edu/utk_infosciepubs

Part of the Library and Information Science Commons

Recommended Citation

Frank, Rebecca D., "Repository Staff Perspectives on the Benefits of Trustworthy Digital Repository Certification" (2023). School of Information Sciences -- Faculty Publications and Other Works. https://trace.tennessee.edu/utk_infosciepubs/475

This Conference Proceeding is brought to you for free and open access by the School of Information Sciences at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in School of Information Sciences -- Faculty Publications and Other Works by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.
REPOSITORY STAFF PERSPECTIVES ON THE BENEFITS OF TRUSTWORTHY DIGITAL REPOSITORY CERTIFICATION

Rebecca D. Frank
School of Information Sciences
University of Tennessee
USA
rfrank7@utk.edu
https://orcid.org/0000-0003-2064-5140

Abstract - This paper reports on the results from a qualitative study that asks whether and how staff members from TRAC certified repositories find value in the audit and certification process. While some interviewees found certification valuable, others argued that the costs outweighed the benefits or expressed ambivalence towards certification. Findings indicate that TRAC certification offered both internal and external benefits, such as improved documentation, accountability, transparency, communication, and standards, but there were concerns about high costs, implementation problems, and lack of objective evaluation criteria.

Keywords – Digital Preservation, Trustworthy Digital Repositories, ISO 16363, TRAC, Repository Assessment

Conference Topics – From Theory to Practice; Sustainability: Real and Imagined.

I. INTRODUCTION

Trustworthy Digital Repositories (TDRs) are organizations that are entrusted with the care and preservation of unique and valuable digital information. From research data, to government records, to cultural heritage information, these repositories ensure the longevity and accessibility of information on a global scale, e.g., [1].

Certification processes have been developed to ensure that the organizations entrusted with this valuable information are indeed able to carry out the work of long-term preservation. Audits carried out by external bodies administer and enforce these certification systems in order to provide assurance to stakeholders that the repositories are trustworthy.

The Trustworthy Digital Repositories: Audit and Certification (TRAC) process, which was strongly influenced by the ISO 16363 standard, is one such certification system [2], [3]. This process is a time-consuming and expensive undertaking for a digital repository, and can result in certification as trustworthy by a team of auditors managed by the Center for Research Libraries [3].

The earliest TRAC certification was issued in 2011, and the most recent in 2015, with an update issued in 2018 for one repository (i.e., CLOCKSS). The staff members of those repositories have therefore had time to reflect on the value proposition of TRAC certification. This paper, which is based on interviews with staff members from all six TRAC certified repositories, asks the following research questions:

- Do staff members from TRAC certified repositories find certification to be valuable?
- How do staff members from TRAC certified repositories characterize the value and/or benefits of TRAC certification?

My findings indicate that while many staff members from TRAC certified repositories find the
audit and certification process to be valuable, and described concrete internal and external benefits, others described the process as more expensive than valuable, and some expressed ambivalence about TRAC certification.

Despite the amount of time that has passed since these TRAC certifications, the ISO 16363 standard, which formed the foundation for the process, was approved in 2012 and was reviewed and confirmed in 2023. This means that current TDR certification processes that rely on, or are influenced by, ISO 16363 are using the same standard as the participants in this research.

II. BACKGROUND

A. Trustworthy Digital Repositories & TRAC Certification

Trust is a central concept in digital preservation [4]–[6]. As early as 1996, members of the digital preservation community identified the need for a mechanism to ensure the trustworthiness of organizations entrusted with the care of unique and valuable digital information [6]. In the nearly 30 years since the Garrett and Waters report, several systems for the audit and certification of digital repositories have emerged, including TRAC, CoreTrustSeal, and nestor e.g., [2], [7], [8].

The TRAC system is based on the ISO 16363 standard, Audit and Certification of Trustworthy Digital Repositories [2]. This certification process is based on the Open Archival Information Systems (OAIS) Model [9], and repository certifications based on this standard have been administered by the Center for Research Libraries (CRL) and the Primary Trustworthy Digital Repository Authorisation Body (PTAB) [10]–[16].

The TDR certification process administered by CRL, TRAC, actively conducted audits from 2011 through 2015 and maintained the certifications awarded through those audits until at least 2018 [17]. The general process for TRAC certification involved repository staff members preparing documentation for review by a team of CRL auditors, followed by a site visit from a small group of auditors who would conduct interviews and inspections in order to assess the veracity of repository documentation [18]. A final determination would be made, and a report prepared for the repository with the findings from the audit team [11]–[16]. The TRAC certification system is the focus of this paper.

TDR certification, including TRAC as well as other systems such as CoreTrustSeal and nestor, is a phenomenon in need of further interrogation. In recent years, scholars such as Maemura, Moles, & Becker have argued that frameworks for repository assessment have not been sufficiently examined [19]. Scholarship about TDR certification has tended to focus on individual reports from organizations that engaged with certification in formal and informal ways e.g., [20]–[24]. Other publications have focused on the development and maintenance of the certification systems e.g., [25]–[29]. There is a need for research that takes a step back from development processes and individual implementations of certification systems to interrogate the value of TDR certification.

B. Benefits and/or Value of TDR Certification

Scholars who have examined the value of TDR certification such as Donaldson have focused on questions about the longevity of digital information in certified repositories, and how certified repositories present this information on their websites [30], [31]. Research has also developed a taxonomy that can be used to address questions about the societal impact of TDRs [32]. A 2018 iPres paper examined the benefits of certification in terms of the return on investment for a particular repository for both Data Seal of Approval and nestor certifications and found that stakeholder confidence, transparent documentation, and process improvement were the most important benefits for their organization [33].

Repositories that have achieved TRAC certification have written about the experience, presenting their certification as a positive development to repository stakeholders e.g., [34], [35], [24]. While much can be learned from this literature, it is unlikely that an organization would be critical of the system in a publication designed to promote their certification. Individuals and organizations involved in the creation of TDR systems have also published informative literature about those systems [25]–[28], [36]. The goal of this category of literature is often promotion of the certification systems, and therefore also has a particular point of view that is unlikely to be critical of TDR certification.

This paper builds upon the scholarship described above to ask whether and how staff members from...
TRAC certified repositories find value in the audit and certification process.

III. RESEARCH METHODS

This paper is part of a larger research project whose goal is to understand risk for long-term preservation in the context of TRAC certification. The project involves interviews with standard developers, auditors, and staff members of TRAC certified repositories. In this paper, I report on the results of 21 interviews with repository staff members from repositories that have received TRAC certification. More information about the research methods, including data collection instruments and the code set used for analysis, is available Open Access at http://hdl.handle.net/2027.42/147539 [37].

A. Data Collection

At the time of data collection in 2016, there were six repositories with TRAC certification: Canadiana.org, Chronopolis, CLOCKSS, HathiTrust, Portico, and Scholars Portal. In-depth, semi-structured interviews were conducted with staff members from all six certified repositories, across three functional areas: repository administration/management, IT, and digital preservation. Previous research has demonstrated that the work of digital preservation involves collaboration across these areas [38].

The interviews, which lasted one to two hours, asked participants to discuss their experiences with the TRAC certification process, and to identify and discuss potential sources of risk for TDRs. Included in the interviews were questions about the cost, benefits, and value of TRAC certification. Audio recordings of the interviews were transcribed for analysis.

B. Data Analysis

Interview transcripts were coded using NVivo. For the first round of analysis, I used a combination of descriptive, analytic, and thematic codes. The code set consisted of codes addressing potential sources of risk, factors that influence the social construction of risk, the TRAC audit process, and attitudes about TDR certification. Working together with another coder to achieve an acceptable level of interrater reliability, we reached a Scott's pi of 0.711 for the subset of interviews with repository staff members [39], [40].

Secondary analysis was conducted by a single researcher, focusing on the topics of cost, benefit, and value of TRAC certification, and attitudes about TRAC certification.

IV. FINDINGS

Findings from this research indicate that the value proposition of TRAC certification is still an open question. While some interviewees described TRAC certification as valuable, others argued that the costs outweighed the benefits. Some also expressed ambivalence about the value of certification. I have organized the findings into four sections based on my analysis: (A) internal benefits, (B) external benefits, (C) arguments that the cost outweighs the benefits, and (D) ambivalence about the benefits of TRAC certification.

A. Internal Benefits of TRAC Certification

TRAC certification was described as valuable for internal repository processes by 12 of the 21 interviewees included in this study.

When asked about the value of certification, interviewees explained that the audit process was valuable because: (1) it forced them to document their policies and practices; (2) the act of creating this documentation enabled them to develop a better understanding of their organization, and to establish a shared understanding of repository policies and practices across the entire organization; and (3) that the review of the documentation by external auditors created an added layer of accountability that ensured a higher quality of documentation than they would otherwise have produced.

The TRAC audit process requires that repositories provide extensive documentation of their policies and processes [3]. Repository Staff 03, 04, 07, 17, and 18 all explained that rather than providing existing documentation to the auditors, their organizations instead had to create current, up-to-date documentation for the purpose of the audit. For example:

“Going through the audit there were a lot of policies you have to have, and we sort of assumed we had them [but] we just didn't have them written down. Going through them we realized in a lot of cases we actually didn't have them.”

(Repository Staff 07)
Similarly, Repository Staff 18 explained that the audit required his organization to formalize internal processes which were not previously documented:

“I think on the technical side, some of it was what we had. On the practice side, I think it was a good exercise ‘cause it forced us to formalize some of these processes that we had done. But we had been doing it internally, but we hadn’t actually said, ‘Okay, well let’s write down a step-by-step guide on how to do this.’ And I think it was useful for us to internally self-organize the archive a little bit.” (Repository Staff 18)

The audit process created an incentive for the organization to create new documentation. For some this was a matter of articulating existing policies more clearly or updating older documentation. For others it meant that repositories had to create policies that did not previously exist. In some cases, the process of creating documentation revealed gaps that were previously unknown to repository staff members:

“[W]e used the same high-level classification of the threats, we certainly identified a lot of things at the operational level where we were not doing as good a job as we should have been. That was a big part of the value of the audit, was that it forced us to actually write down what the processes we were doing were supposed to be, and reviewing whether what was actually happening matched what was supposed to happen. And in many cases it didn’t.” (Repository Staff 13)

Whether they had to create new documentation for the TRAC audit or not, the act of gathering the required information into one coherent set of documents for auditors to review was described as beneficial. This activity created opportunities to share information across different functional areas within a repository, ensuring that the entire staff had a shared understanding of the mission, policies, and practices of the organization:

“I think that having a third party do the audit is much better because you can cheat a lot, inadvertently, when you're doing the self-audit. Just sort of say, “Oh yeah we've got that covered,” without thinking it through. When you actually have to explain to a third party how you've got it covered, that's when you realize that maybe you don't.” (Repository Staff 07)

My findings indicate that the TRAC certification process led to internal benefits for certified repositories, including improved understanding of repository policies and practices, increased...
accountability because of the external auditors, and incentivized the creation of new documentation and formalization of internal processes.

B. External Benefits of TRAC Certification

Interviewees also discussed external benefits from TRAC certification. When discussing these benefits, repository staff members focused on what certification could help them communicate to outside parties, and the role that it allowed their repository to play in the digital preservation community. Nine of the 21 interviewees described TRAC certification as valuable specifically because it (1) improved the transparency of their organization; (2) facilitated communication with repository stakeholders; (3) gave them a competitive advantage in the recruitment of partners, sponsors, and/or funding; and/or (4) gave them an opportunity to be early adopters and establish standards for digital repositories.

Transparency is a central tenet of repository certification [41]. For example, Repository Staff 07 explained that the act of demonstrating trustworthiness by providing information about policies and practices improved his repository’s transparency overall, and that the organization was more proactive about making this information publicly available after certification:

“I also think that there's ongoing value to having that kind of third party oversight in a formal way. But I also think that there is enough oversight now, and there's a lot more transparency on our part just in terms of us being proactive about publishing, and announcing these changes that we make over time, that I'm not as concerned about it.” (Repository Staff 07)

Another benefit of TRAC certification was the fact that certification was seen as communicating something important to repository stakeholders. For some, the goal of TRAC certification was to help stakeholders understand the capabilities of their repository. Both the certification itself, as well as the documentation that repository staff members prepared for the auditors were described as contributing to this benefit.

Repository Staff 07 described certification as a way to establish credibility with external stakeholders: “The reason for doing TRAC certification was to establish credibility in the area and we've done that.” This interviewee went on to explain that they would only maintain certification if the organization could articulate a clear business reason for doing so: “because we’re quite a small organization and because there’s a significant investment of resources, we would certainly be open to doing it, it’s just there would have to be an articulable business reason for doing it” (Repository Staff 07).

In addition to establishing credibility, TRAC certification was also described as something that provided reassurance to stakeholders, “No one has ever proactively asked for it, but when you mention, when I mention it, they shake their heads as though they are reassured in some vague, hard to define way” (Repository Staff 12).

TRAC certification was described as a way to gain a competitive advantage by some interviewees. For repositories with active dues-paying members, for example, certification was viewed as a way to differentiate their organization from others and demonstrate their value. Repository Staff 11 said that the certification helped to recruit members: “[]It has been useful for us to be able to say that we are certified. It’s been useful to be able to say that to libraries and to [partners]. In terms of really practical areas, one of the things we've found is that sort of unexpectedly it brought some new [partners] to us.”

Other repository staff members framed this benefit not as a way to recruit or maintain partners specifically, but rather as a necessary credential to maintain an overall competitive advantage. Repository Staff 13 was confident that his repository would lose business if they did not become TRAC certified: “It was a competitive threat ... Without it [repository] would have lost business.” (Repository Staff 13)

The repositories included in this study were early adopters of repository certification. This was explicitly described as a benefit. Interviewees explained that it was important for their organizations to contribute to the establishment of standards in digital preservation by stepping up to go through this new audit process:

“It also seemed, to me and the team I think, important for us...
These findings demonstrate the ways in which interviewees described external benefits of TRAC certification that focused on what the certification could communicate to external stakeholders, and the role it allowed them to play in the digital preservation community broadly.

C. The Cost of TRAC Certification Outweighs the Benefits

In contrast, six of the 21 interviewees argued that TRAC certification was not valuable for their repository because: (1) the high cost of certification outweighed the benefits; (2) TRAC is not well-known enough to be meaningful; (3) they found problems with the way that certification was implemented.

TRAC certification was described by all the interviewees in this research as very expensive, both in terms of money as well as time that staff members had to spend preparing documentation for the auditors. Some were skeptical about whether these costs outweighed any benefits that they received from certification: “I doubt that the benefits outweigh the costs. I’m sorry to say that. It is not clear to me that the benefits are worth the costs” (Repository Staff 08).

Similarly, Repository Staff 13 said that the TRAC audit process was both costly and disruptive for his repository: “I think there are really big issues about how expensive and disruptive the process is, relative to the benefits that you gain from it. Because there clearly are benefits, but the costs and the disruption are very large” (Repository Staff 13). This interviewee went on to explain that he believed that his organization could have found less costly ways to get the benefits from certification, but that they felt that certification was necessary for financial reasons, “we were under significant competitive pressure. If it hadn’t been for that, we could have got most of these benefits at much lower cost by a more gradual approach, rather than going all the way to TRAC in one go” (Repository Staff 13).

Repository Staff 04 and 11 both felt that the costs of TRAC certification would be barrier for future adoption. Repository Staff 04 argued that cost would need to be lowered substantially for certification to be viable, because the process was prohibitively expensive for his repository. And Repository Staff 11 said that it would take a significant amount of external pressure from stakeholders to go through another audit for recertification: “Honestly, it was such a pain in the butt I am not anxious to do it again. If we started getting pressure from CRL or our libraries or our publishers, then I suspect we would. Without that, my instinct is to coast, actually. It was so much work that, man, we’d have to have a good reason to do it again.”

Repository Staff 13 explained that he would steer any organizations affiliated with his repository away from TRAC certification because of the costs: “There’s no need for any of the other [affiliated] archives at the moment to get certified, and if there was, I wouldn’t recommend that they get 16363, because of the resource implications of trying to do it.”

While the opportunity to be an early adopter was described as a benefit of certification by repository staff members in this research, the relative newness of certification was also seen as a drawback. For Repository Staff 20, a major drawback of certification was that it was relatively unknown and so would not necessarily communicate effectively about his repository’s trustworthiness to others because they would not know what it meant to be TRAC certified: “nobody in [country] had been certified as a Trusted Digital Repository before. So, it was more like not even the process itself but the fact that it was kind of an unknown thing” (Repository Staff 20).

There were several issues raised about the way that TRAC certification was implemented. Interviewees argued that the requirements of TRAC certification were not stringent enough, and that the OAIS model on which TRAC certification was based failed to address the realities of managing a digital repository: “for us the TRAC certification was particularly tricky because TRAC is totally based on OAIS, which totally does not understand a number of aspects of running real world repositories” (Repository Staff 13).
Repository Staff 02 explained that there are no minimum thresholds in TRAC and that repositories could become certified with lots of caveats for sub-optimal policies/practices. Indeed, other research has found that repositories were able to become TRAC certified without fully meeting the requirements outlined in the checklist [18], [42].

Repository Staff 08 expressed dissatisfaction with the auditors. She argued that they were less knowledgeable about digital preservation than the staff of her own repository, which made her doubtful about whether the certification itself held meaning:

“I guess one of my take-aways from the TRAC audit at [repository], and this is my own personal opinion - ... Take it for what it’s worth, when I weigh the level of expertise of the operational team at [repository] against the amount of time and effort put into the documentation used by auditors who in my opinion, please forgive me, were significantly less expert, it made me concerned about the value of the outcome.” (Repository Staff 08)

Six out of 21 interviewees in this study discussed the costs or drawbacks of TRAC certification. They argued that the costs of TRAC certification outweighed the benefits, that TRAC was not well-known enough to be meaningful, and described problems with the way certification was implemented. All interviewees described TRAC certification as very expensive, both in terms of money and time spent preparing documentation for auditors. Some interviewees felt that the costs of certification would be a barrier for future adoption.

D. Ambivalence about the Value of TRAC Certification

In contrast to Sections A, B, and C above in which interviewees argued for or against the value of certification, some interviewees were ambivalent about TRAC. In particular, interviewees were skeptical about the usefulness of the audit outcomes. Interviewees argued that the lack of objective evaluation criteria meant that audit scores were not meaningful and therefore could not be used, for example, to compare their organization against others.

Repository Staff 16 explained that rather than evaluating repositories against an objective set of criteria, the process was designed to assess whether each individual repository was in fact operating in accordance with their own policies: “[T]hey certify that you do what you say you do. They don’t certify that you do something good. Which is a little bit of a vague. So how good you are is what you decide to document and what you decide the processes to be” (Repository Staff 16).

Repository Staff 11 also discussed the flexibility of the TRAC requirements. This interviewee explained that the flexibility was frustrating because it meant that the scores issued for each repository were essentially meaningless and could not be compared against one another:

“That’s one of the interesting things about TRAC, right, is that level of flexibility. It’s also sort of one of the frustrating things about it, too. Because, you know, grades aren’t equal. One institution’s score doesn’t mean they’re providing the same level of preservation as another institution’s score, because you’re evaluating the institution against what the institution said it would do, not against some yardstick.” (Repository Staff 11)

For these interviewees, the value of TRAC certification was an open question, because the scores themselves were viewed as lacking meaningful information about how each repository compared with others. This is particularly interesting in light of the findings from section B above, in which interviewees argued that a key benefit of TRAC certification was that the results facilitated communication with repository stakeholders and conferred a competitive advantage on certified repositories.

Ambivalence about TRAC focused on what the certification could not communicate externally about certified repositories. Staff members from certified repositories believed that it should differentiate their organizations from others by demonstrating their trustworthiness and compliance with best practices. They were frustrated to learn that the results of the process could not be used to make direct comparisons, and that repositories with vastly different practices could receive similar scores.
The flexibility of the TRAC requirements was frustrating for interviewees, as they believed that this meant that the scores issued for each repository were essentially meaningless and could not be compared against one another.

V. DISCUSSION & CONCLUSION

This study highlights the complex and varied perspectives on the value of TRAC certification for digital repositories and the need to continue to examine how certified repositories view the value and benefits of the process over time. My findings indicate that staff members of TRAC certified repositories understood certification to have both internal and external benefits for repositories, including improved transparency, communication with stakeholders, and a competitive advantage in recruitment of partners, sponsors, and funding. However, some interviewees argued that the high cost of certification outweighed the benefits, that TRAC was not well-known enough to be meaningful, and that there were problems with the way certification was implemented. Additionally, some interviewees expressed ambivalence about the value of TRAC certification, arguing that the lack of objective evaluation criteria meant that the audit scores were not meaningful.

This aligns with findings from my previous research in which I found that the highly flexible certification criteria, which are intended to allow the system to be applicable across a broad array of repository types, have been used by repositories to justify sub-optimal preservation practices e.g., [18], [42]. In this paper I argue that this flexibility, which I have characterized elsewhere as a potential source of risk for both digital repositories and the long-term preservation of the digital information they contain, also detracts from the value of certification for some stakeholders.

This study complements previous research about the value of TDR certification. For example, Donaldson has carried out research which seeks to understand whether repositories with TDR certification have better long-term outcomes, in order to understand the impact of certification [32]. Notably, my findings show that despite the benefits listed here, participants did not say that the information in their repositories was more secure or better preserved after completing a TRAC audit. Also absent were arguments that their repositories were more trustworthy or better able to preserve information long term as a result of going through the TRAC audit process. Rather, the benefits centered on aligning the expectations of internal and external stakeholders, and of improving transparency and communication in order to remain competitive.

As discussed in Section II B above, much of what is known about TDR certification has been produced by those involved in the process in some way – developers of certification systems, and repositories that have achieved certification. This paper provides a new perspective, investigating the value of TRAC certification through empirical research. Even so, participants in this study may still have been motivated by a desire to promote the certification system. Achieving TRAC certification was a costly endeavor and phenomena such as escalation of commitment and/or sunk cost bias may have been present in this study [43], [44].

Future research, which considers both the repository outcomes as well as the attitudes and beliefs of repository staff members has the potential to produce a more complete picture of the value of this relatively new phenomenon. Additionally, as more time passes, repository stakeholders may be willing and/or able to reflect on their experiences with TDR certification in different ways.

Finally, TRAC is one of several TDR certification systems that are active today. While some of the criticism about TRAC certification focused on the requirements themselves, much centered on the particular implementation of TRAC certification as administered by CRL. More recent audits have been conducted by a different organization (i.e., PTAB), and future research should investigate this new implementation of the ISO 16363 standard [10].

VI. ACKNOWLEDGMENTS

I would like to thank Megh Marathe and Carl Haynes for their assistance with data analysis. Thank you also to Elizabeth Yakel, Ph.D. for support and thoughtful feedback across the lifecycle of the project. This research was funded in part by a University of Michigan Rackham Graduate Student Research Grant.

REFERENCES


