RESEARCH ARTICLE

JASIST WILEY

An empirical examination of data reuser trust in a digital repository

Elizabeth Yakel[1] | **Ixchel M. Faniel**[2] | **Lionel P. Robert Jr**[1]

[1]University of Michigan School of Information, Ann Arbor, Michigan, USA

[2]OCLC, Dublin, Ohio, USA

Correspondence
Elizabeth Yakel, University of Michigan School of Information, 105 South State Street, Ann Arbor, MI 48109, USA.
Email: yakel@umich.edu

Abstract

Most studies of trusted digital repositories have focused on the internal factors delineated in the Open Archival Information System (OAIS) Reference Model—organizational structure, technical infrastructure, and policies, procedures, and processes. Typically, these factors are used during an audit and certification process to demonstrate a repository can be trusted. The factors influencing a repository's designated community of users to trust it remains largely unexplored. This article proposes and tests a model of trust in a data repository and the influence trust has on users' intention to continue using it. Based on analysis of 245 surveys from quantitative social scientists who published research based on the holdings of one data repository, findings show three factors are positively related to data reuser trust—integrity, identification, and structural assurance. In turn, trust and performance expectancy are positively related to data reusers' intentions to return to the repository for more data. As one of the first studies of its kind, it shows the conceptualization of trusted digital repositories needs to go beyond high-level definitions and simple application of the OAIS standard. Trust needs to encompass the complex trust relationship between designated communities of users that the repositories are being built to serve.

1 | INTRODUCTION

Trust is a central characteristic of digital curation, preservation and continued (re)use. In the 2002 report, *Trusted Digital Repositories*, the authors identify three sites of trust surrounding digital repositories: user trust in repositories, trust in third-party providers, and user trust in data (Research Libraries Group-OCLC, 2002, p. 9). Over 20 years later, we still have limited research on user trust in repositories (Yakel et al., 2013; Yoon, 2016). This matters now more than ever as trust in all types of organizations has diminished, including archives and digital repositories (Bak, 2016; Edelman, 2022; Saad, 2023; Pew Research Center, 2020; Price & Smith, 2011). At the same time, funders, publishers, and disciplinary communities are mandating or strongly encouraging data sharing and thus a concurrent increased reliance on digital repositories to manage those data and the sharing process (Nature, n.d.; Cramer, 2022; National Institutes of Health, 2023; White House Office of Science and Technology Policy, 2023). A digital repository is formally defined as "an Archive consisting of an organization, which may be part of a larger organization of people and systems that has accepted the responsibility to preserve

information and make it available for a Designated Community" (Consultative Committee for Space Data Systems, 2011, p. 1). Digital repositories do more than store data; they are knowledge infrastructures—a web of interconnections among people, artifacts, and institutions to create, share, and curate knowledge (Edwards, 2010). As such, repositories have fostered the development of disciplines, enhanced and ensured the transmission of research methods, and set data standards (Ribes & Bowker, 2009; Shankar et al., 2016). Given this crucial relationship between repositories and their users, it is surprising how little research has focused on trust between data reusers and repositories.

"[T]rust can be viewed as something (e.g., acceptance, approval, confidence, or respect) that is sought or something that can be bestowed" (Prieto, 2009, p. 593). The *Open Archival Information System (OAIS) Reference Model* is the standard for developing trusted digital repositories that preserve access to digital content over the long term Consultative Committee for Space Data Systems. It provides a common framework to demonstrate that repositories can be trusted by stakeholders, such as funders, certification bodies, parent organizations, and users. *Audit and Certification of Trustworthy Digital Repositories* (Consultative Committee for Space Data Systems, 2011) is the parallel assessment standard that assures repositories have sought out and met the OAIS standard. To date, most studies of trusted digital repositories have focused on repositories' internal operations, such as organizational structure, technical infrastructure, and policies, procedures, and processes, as set out in the OAIS standard from the perspective of certification bodies (e.g., Jantz & Giarlo, 2007; Schultz & Gore, 2010). This has left the factors that influence users to bestow trust in digital repositories more theoretical than empirical (Prieto, 2009) as well as left the connection between trust and continued use of a digital repository unexplored.

Although it has been 15 years since Prieto's (2009) call for additional empirical research investigating user perceptions of trust in digital repositories, this has not been heeded. Drawing from the literature in archives and trusted digital repositories as well as the literature on trust in organizations and information systems, we propose and test a model of trust in a digital repository specifically for data (hereafter data repository) that we believe is critical for continued use. More specifically, we study data reusers and examine the factors that influence their trust in a data repository and the influence trust has on their intention to continue using it. As the first study to quantitatively examine trust in data repositories, we propose and then test a relationship between trust, the trustworthiness factors that support it, and

continuance intention. Our study is based on an analysis of 245 survey respondents who published research based on the holdings of one data repository. We found three trustworthiness factors were positively related to data reuser trust—integrity, identification, and structural assurance. In turn, trust and performance expectancy were positively related to data reuser intention to return to a repository for more data.

2 | OUR APPROACH TO THE PROBLEM OF TRUST IN DATA REPOSITORIES

Data reuser trust is the central factor in our model of trust in data repositories. One we believe is critical to the success of data repositories. Our model presents a complex formulation of data reuser trust examining both the trustworthiness factors leading to trust and the outcomes of trust. This delineation aligns with other models of trust. For example, in their proposed model of dyadic trust in organizations, Mayer et al. (1995) make distinctions between trust, factors that contribute to it, and outcomes that result from it.

As background, trust is an ill-defined concept in the archival and digital curation literature. The 2002 RLG-OCLC report relies on the *Merriam Webster Dictionary* definition of trust as "assured reliance on the character, ability, strength, or truth of someone or something... one in which confidence is placed... a charge or duty imposed in faith or confidence or as a condition of some relationship... something committed or entrusted to one to be used or cared for in the interest of another" (Research Libraries Group-OCLC, 2002, p. 8).

While this report is referenced in OAIS, trust is not defined or operationalized in OAIS (Bak, 2016). Even articles about trust in repositories often fail to define it (Oliver et al., 2011; Speck, 2010) or use the dictionary definition (Price & Smith, 2011). Recently, drawing on the organization and information systems literature, several scholars have attempted to elucidate the definition of trust in digital curation and emphasized the central relationship of trust between a designated community and a repository or data (Prieto, 2009; Yoon, 2014; Yoon & Lee, 2019). Prieto draws on the organization and information systems literature to posit a user-centric conceptualization of trust in digital repositories and argues "User communities are the most valuable component in ensuring a digital repository's trustworthiness. It is important, then, to study their perceptions of trust as factors critical to the success of digital repositories" (Prieto, 2009, p. 603). Yoon (2014) employs the organization and information systems literature to help explain

how data reusers conceptualize trust in data repositories. She also uses this literature to operationalize factors in her quantitative investigation of data reusers trust in data (Yoon & Lee, 2019). Both Yoon and Prieto apply different concepts and propose different operationalizations of trust from this literature; however, they both center repository stakeholders in their models of trust.

Trust is "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (Mayer et al., 1995, p. 712). An important feature of trust is risk, which is essential to behavioral manifestations of trust (i.e., the actions one takes) (Mayer et al., 1995). In the information systems literature that action is described as the intention to use the information system under study. More specifically, Meeßen et al. (2020) propose that "if users are willing to depend on and be vulnerable to an MIS, they should have a greater intention to use it" (p. 11). The literature on trustworthiness factors and trust is extensive and includes research examining trust among staff in organizations, trust of staff in organizations, as well as people's trust in the organizations themselves (e.g., Mayer et al., 1995; Oliver et al., 2011; Pirson & Malhotra, 2011). A stream of literature particularly pertinent to our study is people's trust in information systems given the perceived risk associated with the online exchange of information, goods, and services. Empirical studies support the assertion that trust in these types of information systems has been found to influence an individual's intention to use them (e.g., Akter et al., 2011; Gefen et al., 2003; Gefen & Straub, 2004; Komiak & Benbasat, 2006; Li et al., 2008; McKnight et al., 2002; Meeßen et al., 2020). Studies have adapted and extended models of trust in the context of other trustors and trustees, such as stakeholder trust in organizations and consumer trust in e-retailers (Kimery & McCord, 2002; Pirson & Malhotra, 2011). Consequently, the results depict a complex and sometimes contradictory landscape (Akter et al., 2011; Gefen et al., 2003; Meeßen et al., 2020).

In some information systems studies, trustworthiness factors are treated as measures of trust that influence intention (e.g., Gefen & Straub, 2004; Hallikainen & Laukkanen, 2021; Wang & Benbasat, 2008). In others, they are modeled as underlying constructs of trustworthiness and trustworthiness is treated as a second-order construct that influences trust (e.g., Akter et al., 2011; Serva et al., 2005; Xie et al., 2020). Additionally, there are studies where the trustworthiness factors are modeled as a grouping of first-order constructs and each factor's relationship with trust is tested individually (e.g., Komiak &

Benbasat, 2008; Lee & Turban, 2001; Serva et al., 2005). Finding that both first- and second-order models of trustworthiness have acceptable fits, Serva et al. (2005) conclude the choice between the two depends on the research objectives.

We developed our model of data reuser trust based on existing research in the organizational and information systems literature. As these studies all note the importance of contextualization in the rationale concerning which factors to choose, we used the archival and digital curation literature to ground our concepts in the context of data repositories. Furthermore, given the lack of research on trust in data repositories, we tested the relationship between each trustworthiness factor and trust individually to provide a more nuanced understanding about the role trustworthiness factors have in developing data reuser trust in data repositories and how in turn that trust leads to the trusting action continuance intention.

3 | A MODEL OF DATA REUSER TRUST IN A DIGITAL REPOSITORY

Drawing from the organization and information systems literature, we define trust in a data repository as the willingness of the trustor (i.e., data reuser) to accept vulnerability in an online exchange of data based on positive expectations of the trustee's (i.e., data repository) future behaviors (adapted from Mayer et al., 1995, p. 712). We begin this section with a discussion of the five trustworthiness factors—integrity, benevolence, transparency, identification, and structural assurance which we hypothesize lead to trust. We then discuss the factors related to continuance intention to use a data repository—trust, performance expectancy, and social factors. Finally, we present several moderating factors. We hypothesize that reuse reliance and data scarcity moderate the relationship between trust and continuance intention and reuser experience moderates the relationship between social factors and continuance intention. This interdisciplinary exploration grounds our investigation and elucidates our decisions for including certain concepts and excluding others and provides rationale for our hypotheses.

3.1 | Trustworthiness factors

Integrity encompasses consistency of actions, sense of justice, and congruence between words and actions (Mayer et al., 1995). Organizational and information systems studies have described it as an adherence to principles (e.g., Akter et al., 2011; McKnight et al., 2002;

Pirson & Malhotra, 2011; Wang & Benbasat, 2008). Evaluating the integrity of a data repository is critical for demonstrating that it can be trusted by the community it supports (Consultative Committee for Space Data Systems 2011; Prieto, 2009; Ross & McHugh, 2006). In other words, it needs to be held to the highest standards by reporting lapses and the steps taken to address them as well as setting and meeting expectations and service levels. Garrett and Waters (1996) cite integrity as a key trustworthiness factor for digital archives, they continue "[u]sers of archived information in electronic form and of archival services relating to that information need to have assurance that a digital archives is what it says that it is and that the information stored there is safe for the long term" (Garrett & Waters, 1996, p. 23). Qualitative studies lend support to these ideas, showing users' trust is related to their perceptions that data are valid and the data repository is not misleading or deceiving (Yoon, 2013, 2014). For these reasons, we hypothesize that data reusers are more likely to trust a data repository when they believe it demonstrates integrity.

> **Hypothesis 1.** Integrity is positively related to data reuser trust in a data repository.

Benevolence speaks to intentions and motives of the trustee (Mayer et al., 1995). In the organization and information systems literature, benevolence is characterized as acting in the interest of another, being willing to serve, not being opportunistic or manipulative (e.g., Akter et al., 2011; McKnight et al., 2002; Pirson & Malhotra, 2011; Wang & Benbasat, 2008). Benevolence describes the extent a trustee "exhibits goodwill toward the trustor and has concern for the trustor's well-being" (Pirson & Malhotra, 2011, p. 1090). Although data repositories often exist in academic, research, and cultural heritage settings where profits are not the primary motive, they have different sustainability models. Rather than provide free, open access to data, some are membership driven. Despite these differences, a trusted repository is expected to care about the mission "to provide reliable, long-term access to managed digital resources to its designated community, now and into the future" (Research Libraries Group-OCLC, 2002, p. i and 5). This means that it will utilize the best and most reliable techniques and services for curating and making data accessible. These goodwill gestures demonstrate a data repository's benevolence. We hypothesize that a data repository perceived to be benevolent is more likely to be trusted by data reusers.

> **Hypothesis 2.** Benevolence is positively related to data reuser trust in a data repository.

Transparency signifies the visibility and accessibility of information and measures the extent the trustee gives "access to information regarding organizational behaviors and intentions" (Pirson & Malhotra, 2011, p. 1091). Described in terms of the amount and type of information made visible or accessible, experimental studies have examined how different transparency features for websites and recommender systems influence trust (e.g., Cramer et al., 2008; Grimmelikhuijsen, 2009; Wang & Benbasat, 2008). The results have been mixed. Transparency is, however, a frequently mentioned trust factor in the archival and curation literature. Price and Smith (2011) claim it is part of a dominant discourse around repositories. A repository is expected to subject its design specifications, practices, policies, and procedures to a risk analysis, the results of which help stakeholders make informed decisions (Consultative Committee for Space Data Systems, 2011). "Communicating audit results to the public—transparency—will engender more trust, and additional objective audits, potentially leading towards certification, will promote further trust in the repository and the system that supports it" (Consultative Committee for Space Data Systems, 2012, p. 19). More specifically, its users must be made aware of the methods used to evaluate preservation choices (e.g., tools used to assess authenticity and integrity of digital records over time) to determine whether the methods meet their needs (Adams et al., 1992; Corrado, 2019). "Operational transparency is yet another way of inviting public scrutiny and strengthening the bond with archives" (Speck, 2010, p. 48). Users also need to know what actions have been taken on digital content over time (Beagrie, 2013; Carlson & Anderson, 2007). Although transparency is a key principle used when applying audit and certification criteria for trustworthy repositories, to our knowledge only two qualitative studies have examined whether a repository's transparency is related to trust and the results are mixed (e.g., Yakel et al., 2013; Yoon, 2014). We hypothesize that data reusers are more likely to trust data repositories they believe are being transparent.

> **Hypothesis 3.** Transparency is positively related to data reuser trust in a data repository.

Pirson and Malhotra (2011) describe identification as "the understanding and internalization of the interests and intentions of the other party" (p. 1090). Pirson and Malhotra found that identification was a significant trustworthiness dimension for all stakeholders, regardless of the number of interactions with the organization. Alternatively, Lewicki and Bunker (1995) found that

identification-based trust became more important in the mature stages of trust development. Based on the digital curation literature, identification is implicit in the mandate to understand and respond to the changing needs of designated communities (Consultative Committee for Space Data Systems, 2012). That said, responding to a designated community is a difficult to impossible task for repositories, particularly those that have a broad user base, and must be done "with input from actual users and communities" (Bettivia, 2016, p. 8). Prior research shows some repositories use "controlled change to manage fluidity" (Daniels et al., 2012, p. 286) when adapting to their designated communities' changing practices by checking in with data producers as they process data, waiting for their designated communities to reach a consensus, and ensuring changes could be rolled back to prior states. The benefits of this kind of user-focused attention can be seen in the history of social science data archives, which have grown and changed over time, due in part to the methods they use to identify with their designated communities, through education programs, changing collection goals, and outreach to potential data depositors and reusers (Shankar et al., 2016). We hypothesize that data reusers are more likely to trust a data repository that understands disciplinary practice in their scholarly community.

> **Hypothesis 4.** Identification is positively related to data reuser trust in a data repository.

Structural assurance refers to organizational level factors, such as internal and external structures; regulations or safeguards; and legal remedies that are put in place to address problems (Gefen et al., 2003; McKnight et al., 2002; Sitkin & Roth, 1993). Rousseau et al. (1998) describe these kinds of structures as institutional factors that "can act as broad supports for the critical mass of trust that sustains further risk taking and trust behavior" (p. 400). Some literature includes ability which is defined as "that group of skills, competencies, and characteristics that enable a party to have influence within some specific domain" (Mayer et al., 1995, p. 717). Based on our review of previous research in the information systems literature, we have substituted structural assurance for ability since it pertains more to information systems. Studies in information systems have shown trust formation is less about declaring skills and competencies than it is about demonstrating them through experiences and interactions (Wang & Benbasat, 2008). This is particularly true for trusted digital repositories. Organizations should not simply "declare themselves 'OAIS compliant' to

underscore the trustworthiness of their digital repositories" (Consultative Committee for Space Data Systems, 2012, p. 1-1). At the very least, long-term preservation relies on assessing a digital repository "to understand its capabilities, where it stands against potential threats, and any other risks inherent in its systems" (Consultative Committee for Space Data Systems, 2011, p. 3). Positive results from this process can act as a guarantee and provide a means through which users can establish trust in the repository (Ross & McHugh, 2006). Guarantees are one form of structural assurance, but data repositories have others, such as third-party endorsements and repository reputation (Yakel et al., 2013; Yoon, 2014). However, information systems studies show conflicting results (Gefen et al., 2003; Kimery & McCord, 2002; Li et al., 2008; McKnight et al., 2002; Xie et al., 2020) due to the different ways trustworthiness and trust have been conceptualized and measured as well as the institutional context (Li et al., 2008; McKnight et al., 2002). For instance, McKnight et al. (2002) suggest distinguishing between different types of internet experiences (e.g., information, shopping, and advice sites). Considering the conflicting evidence, structural assurance is included in our model as repository reputation has consistently been identified with trust (Bak, 2016; Yakel et al., 2013; Yoon & Lee, 2019) although the results of its effect on trust have been inconsistent. We hypothesize that data reusers are more likely to trust a data repository that provides some type of structural assurance.

> **Hypothesis 5.** Structural assurance is positively related to data reuser trust in a data repository.

3.2 | Continuance intention

An important feature of trust is risk, which is essential to the behavioral manifestation of trust (i.e., the action one takes) (Mayer et al., 1995). In the information systems literature that action is described as the intention to use the information system under study. Usage intention is a key indicator in technology acceptance (e.g., Davis, 1989; Taylor & Todd, 1995; Venkatesh et al., 2003). Both first-time and continued use of technology have been studied in models of trust in technology (e.g., Akter et al., 2011; Hallikainen & Laukkanen, 2021; He et al., 2009). Bhattacherjee (2001) coined the term information system or IS continuance intention to describe "users' intention to continue using" (p. 359), which is of particular interest for a data repository. More specifically, Meeßen et al. (2020) propose that "if users are willing to depend on and

be vulnerable to an MIS, they should have a greater intention to use it" (p. 11). In our study, the trusting action continuance intention signifies a data reuser's intention to continue using a data repository. Studies show individuals' trust in web-based technology is significantly related to their intention to continue using it (e.g., Akter et al., 2011; He et al., 2009). Data repositories offer data through web-based technologies. The archival and digital curation literature has discussed the factors that lead to trust in repositories with the assumption that trust leads to use (Consultative Committee for Space Data Systems, 2011; Prieto, 2009). But to our knowledge there are no studies that have tested that relationship. We hypothesize data reusers' trust in a data repository is related to the intention to continue using it.

> **Hypothesis 6.** Trust is positively related to data reuser continuance intention.

Performance expectancy is "the degree to which an individual believes that using the system will help him or her attain gains in job performance" (Venkatesh et al., 2003, p. 447). The significance of its relationship to intention has been consistent across studies of various users of information systems, such as email, word processing programs, computing resource centers, online banking, and other web-based systems (e.g., Bhattacherjee, 2001; Davis, 1989; Davis et al., 1992; Gefen et al., 2003; Taylor & Todd, 1995). Moreover, relationships between performance expectancy and usage intention have remained significant over time (Davis, 1989; Davis et al., 1992; Venkatesh et al., 2003). Data reusers often do not have the resources to collect their own data or the network of colleagues willing and able to share theirs. Data repositories are meant to address these issues. By providing access to data, repositories create the opportunity to publish more studies, more quickly, which has implications for data reusers' job performance (e.g., tenure, promotion, accelerated research discoveries). Although Curty (2015) found that "perceived effort did not have a significant negative influence on social scientists' intention to reuse data" (Curty, 2015, p. 203) her participants did cite data access and discovery as hinderances to data reuse. Taken together these findings lend credence to including performance expectancy in our study. Therefore, like other technology acceptance studies, we expect data reusers who believe data repositories are capable of helping them in these ways are likely to continue using it.

> **Hypothesis 7.** Performance expectancy is positively related to data reuser continuance intention.

The relationship of social factors (also known as social norms, social influence) to usage intention also has been studied (e.g., Davis, 1989; Taylor & Todd, 1995; Thompson et al., 1991; Venkatesh et al., 2003). Unlike performance expectancy, findings are mixed (e.g., Davis, 1989; Taylor & Todd, 1995). They have been explained as differences in settings, such as mandatory versus voluntary systems use (Barki & Hartwick, 1994; Venkatesh & Davis, 2000) and individual differences, such as age, gender, and experience (e.g., Karahanna et al., 1999; Venkatesh et al., 2003; Venkatesh & Morris, 2000). Social factors describe a data reuser's "internalization of the reference group's subjective culture and specific interpersonal agreements that the individual has made with others in specific social situations" (Triandis, 1980, p. 2010 as quoted in Thompson et al., 1991, p. 126). Studies show data reusers' decisions about whether and how to reuse data are influenced by experts and perceived norms within their communities (Borgman, 2015; Curty, 2015; Faniel et al., 2012). Kriesberg et al. (2013) describe the role mentorship plays in the enculturation of novice data reusers into disciplines through a cognitive apprenticeship process that extends beyond how to reuse data to understanding and applying the ethics and norms for evidence within their disciplines. For these reasons, social factors are hypothesized to positively influence data reusers' decisions to use repositories.

> **Hypothesis 8.** Social factors are positively related to data reuser continuance intention.

3.3 | Moderating factors

In our model of trust in data repositories, we consider three moderating variables—reuse reliance, data scarcity, and reuse experience. Reuse reliance is defined as the percentage of a reuser's research that depends on data collected by others. Rousseau et al. (1998) contend that interdependence is a necessary condition of trust "where the interests of one party cannot be achieved without reliance upon another" (p. 395). They argue that changes in the level of interdependence alters the level of trust. The reasons for reuse reliance vary. Disciplinary culture and norms that support reuse can influence its frequency and acceptability (Borgman, 2015; Curty et al., 2017; Darch, 2018). Data collection is time and resource intensive for some researchers given limited access to graduate students, expensive equipment, and travel. Research interests and approaches require some researchers to supplement their own data collection with data from others (e.g., administrative data collected at the local, state, or federal level). Some need to reuse data to calibrate their

numerical models (Curty et al., 2017; Faniel & Jacobsen, 2010). Others rely on data they perceive are easy to reuse, because the data are highly standardized, not sensitive, and have fewer intellectual property restrictions (e.g., Curty et al., 2017). It is not whether reusers encounter these situations, it is how frequently they encounter them that speaks to their reliance on reuse. Those who have high reuse reliance are likely to learn what repositories they trust. In this case, reuse reliance is expected to strengthen the relationship between trust and continuance intention.

> **Hypothesis 9.** Reuse reliance moderates the relationship between trust and continuance intention.

Data-scarce domains are those that do not have "enough data to pursue the domain's major objectives" (Darch, 2018, p. 166). Data scarcity can also exist for individual researchers who have insufficient data to pursue research objectives. It can be particularly problematic for those pursuing novel areas of study or wanting to test new research methods. Other factors that contribute to data scarcity include difficulty finding data (Faniel et al., 2013) as well as complex de-identification processes and restrictive terms of use that hinder data sharing (Schäfer et al., 2011). Curty (2015) discusses the complexities of deidentification and terms of use in reference to qualitative human-subjects data, but these complexities also are present in quantitative data and fieldwork occurring in vulnerable locations (Frank et al., 2015). In these kinds of data-scarce environments, there are fewer opportunities to engage with and build trust in data repositories. It is expected that data scarcity will weaken the relationship between trust and continuance intention.

> **Hypothesis 10.** Data scarcity moderates the relationship between trust and continuance intention.

Reuse experience is defined as the number of years a researcher has been reusing data. Those with more reuse experience are likely to be more skilled at deciding which data repositories to use. As experts, they likely serve as mentors, shape practice, and influence behavior within their disciplinary communities (Kriesberg et al., 2013; Yoon, 2015). Over time as their expertise grows, their intention to continue using a data repository is less likely to be influenced by social factors.

> **Hypothesis 11.** Reuse experience moderates the relationship between social factors and continuance intention.

4 | METHODS

This study was done in partnership with the Inter-university Consortium for Political and Social Research (ICPSR). ICPSR was selected as the study site for several reasons. First, ICSPR is a leading social science data archive at the forefront of social science data preservation, access, and curation. Second, its longevity (founded in 1962) and the breadth of quantitative and qualitative holdings (250,000 data files) have enabled the institution to engage with multiple disciplinary communities (ICPSR, n.d.). ICPSR recruits data from major studies and contracts with several survey organizations and federal agencies to obtain their data for preservation; funders also mandate data deposit for some projects. Finally, ICPSR tracks data reuse. ICPSR provides data depositors information about the number of downloads their data set receives as well as citations to studies that reuse the data. The citations are listed on the homepage for each data set and in the Bibliography of Data-Related Literature found on the ICPSR website.

4.1 | Survey development

Most survey items were adapted from the literature (see Table 1, for references). We revised questions to contextualize them for a data repository. We used the original 7-point Likert scales ranging from 1 (strongly disagree) to 7 (strongly agree), with the exception of the moderating variables and the demographic questions. The survey was piloted twice. During the first pilot, three cognitive walk-throughs were conducted. Anderson and Gerbing (1991) recommend testing with subject matter experts who are representative of the population. Each walkthrough consisted of a one-on-one session with an individual who reused data. Using the concurrent think-aloud technique, respondents verbalized their thoughts as they answered the survey questions (Groves et al., 2009). This helped clarify several survey items, assess content validity (Straub et al., 2004), and confirm the questions were measuring the constructs as intended (Ryan & Bernard, 2003). The second pilot employed Qualtrics, a web-based survey administration platform. Forty-four social scientists were invited to complete the survey, twenty-seven responded. Data from the second pilot helped determine survey timing and select final survey items.

4.2 | Survey sample

ICSPR's Bibliography of Data-Related Literature was used to identify 8,461 citations to data deposited in ICPSR

TABLE 1 Factor loadings.

Constructs	Items	1	2	3	4	5	6	7	8	9
Benevolence (Pirson & Malhotra, 2011)	ICPSR cares about users	**0.86**	0.47	0.70	0.63	0.50	0.38	0.53	0.68	0.57
	ICPSR listens to my needs	**0.75**	0.30	0.60	0.50	0.41	0.34	0.48	0.56	0.42
	ICPSR does not abuse users	**0.83**	0.40	0.55	0.77	0.43	0.39	0.51	0.68	0.58
Continuance intention (Bhattacherjee, 2001; Kim & Han, 2009)	I will continue my use of ICPSR in the future	0.50	**0.90**	0.52	0.48	0.71	0.43	0.53	0.54	0.71
	I will keep using ICPSR as regularly as I do now	0.38	**0.84**	0.42	0.42	0.64	0.40	0.55	0.46	0.56
	If I could, I would like to discontinue my use of ICPSR	0.30	**0.74**	0.28	0.25	0.48	0.24	0.33	0.28	0.48
	I will continue using ICPSR rather than use any alternatives	Removed								
Identification (Pirson & Malhotra, 2011)	ICPSR understands practice in my discipline	0.60	0.33	**0.81**	0.62	0.42	0.39	0.49	0.54	0.54
	ICPSR responds to users' requests for better contextual information around data	0.57	0.23	**0.66**	0.47	0.35	0.34	0.42	0.54	0.37
	ICPSR provides search and discovery tools that align with practice in my discipline	0.60	0.55	**0.84**	0.52	0.69	0.50	0.58	0.62	0.65
Integrity (Pirson & Malhotra, 2011)	ICPSR does not try to deceive	0.55	0.41	0.48	**0.77**	0.48	0.33	0.51	0.54	0.51
	ICPSR has high ethical standards	0.69	0.38	0.58	**0.87**	0.46	0.45	0.57	0.63	0.64
	ICPSR treats its users with respect	0.75	0.42	0.68	**0.89**	0.48	0.42	0.57	0.66	0.60
Performance expectancy (Venkatesh et al., 2003)	Using ICPSR enables me to accomplish my research more quickly	0.53	0.66	0.62	0.51	**0.88**	0.42	0.58	0.53	0.64
	Using ICPSR improves my research performance	0.48	0.66	0.60	0.49	**0.88**	0.39	0.48	0.49	0.60
	I have found ICPSR useful for my research	0.43	0.66	0.51	0.47	**0.86**	0.45	0.46	0.49	0.63
Social factors (Thompson et al., 1991)	My senior colleagues have been helpful in the use of the ICPSR repository	0.19	0.19	0.21	0.14	0.17	**0.58**	0.36	0.23	0.19
	My mentors are very supportive of the use of the ICPSR repository for my research	0.39	0.36	0.50	0.41	0.42	**0.82**	0.43	0.42	0.46
	In general, my institution has supported my use of the ICPSR repository	0.41	0.41	0.46	0.45	0.44	**0.86**	0.50	0.35	0.44
	I use the ICPSR repository because of the proportion of my peers who use it	Removed								
Structural assurance (Gefen et al., 2003)	I feel confident conducting my research using data from ICPSR because I can contact user support for help	0.51	0.38	0.48	0.40	0.34	0.38	**0.70**	0.46	0.39
	I feel confident conducting my research using data from ICPSR because of its stated commitments to best practices	0.48	0.32	0.47	0.53	0.35	0.47	**0.80**	0.49	0.49
	I feel confident conducting my research using data from ICPSR because it is a well-known, reputable repository	0.48	0.59	0.55	0.57	0.59	0.47	**0.82**	0.54	0.72
	I feel confident conducting my research using data from ICPSR because it has received the Data Seal of Approval	Removed								
Transparency (Pirson & Malhotra, 2011)	ICPSR explains its decisions about data	0.57	0.33	0.59	0.58	0.38	0.32	0.49	**0.79**	0.45
	ICPSR is transparent	0.62	0.53	0.59	0.61	0.53	0.38	0.60	**0.86**	0.58
	ICPSR openly shares all relevant information	0.59	0.39	0.52	0.59	0.41	0.38	0.48	**0.79**	0.45
	ICPSR provides adequate feedback and reporting mechanisms	0.73	0.42	0.63	0.55	0.48	0.34	0.47	**0.74**	0.49

T A B L E 1   (Continued)

Constructs	Items	1	2	3	4	5	6	7	8	9
Trust (Pirson & Malhotra, 2011)	I trust ICPSR	0.65	0.59	0.65	0.71	0.64	0.47	0.64	0.62	**0.87**
	I would recommend ICPSR because I believe it is a stable organization	0.49	0.58	0.56	0.57	0.54	0.39	0.67	0.48	**0.87**
	I would recommend using data from ICPSR to colleagues	0.56	0.70	0.60	0.54	0.69	0.47	0.60	0.54	**0.89**

Note: Extraction method: principal component analysis. Rotation method: Oblimin with Kaiser normalization. Bold indicates the construct on which items loaded most highly.

between 2006 and 2012. The scope was later narrowed to journal articles published between 2009 and 2012. Next, the first authors were identified. In cases where an author appeared two or more times as the first author of a journal article, only the most recent citation was retained. Authors who also were producers of the cited data were eliminated. Next, we searched for the email of the first authors, limiting our searches to 5 min each. This eliminated more names. At the end of this process, the sample size was smaller than planned. To increase the sample size, citations to articles published in 2008 were drawn from the 10 most frequently used data sets deposited in ICPSR. The same method was used to identify first authors and search from their emails.

Only a limited number of true demographic questions were asked—field and academic track. Findings indicate a highly diverse sample. Almost 50% of respondents designated Sociology (31.0%) or Criminology and Criminal Justice (18.0%) as their primary fields, but Public Health (13.5%), Economics (9.8%), and Psychology (4.9%) also were represented. The majority of respondents (83.2%) were tenured or tenure track—full (26.9%), associate (30.6%), or assistant (25.7%) professors. Small percentages of lecturers, graduate students, researchers, and others comprised the remaining respondents. Several questions were also related to reuse: How many years have you been reusing data in your research? (Reuse experience); What percentage of your research relies on data collected by other people? (Reuse reliance); Are there sufficient data available for reuse in your field? (Data scarcity); In your research have you used data from any of the following sources? (General source); and Have you contributed data to ICPSR? (Contributor). Overall, respondents had been reusing data for a considerable length of time, the mean was 13.2 years, although the range spanned 1–46 years. The percentage of respondents reporting that more than 75% of their research relied on reuse was 57.5%. A third (33.1%) reported sufficient data were not available for reuse. Respondents used a variety of data sources, including collecting one's own data (69.4%);

local, state, and federal governments (66.9%); colleagues (60.4%); and repositories other than ICPSR (59.2%). Only 15.9% had contributed data to ICPSR.

4.3 | Survey administration

In the week before survey administration via Qualtrics, the ICPSR director sent an email informing potential survey respondents about the survey and encouraged participation. One week later, personalized emails were sent to the respondents inviting them to complete the survey. Three additional follow-up notices were sent in successive weeks and the survey was closed after 6 weeks. Our final sample size was 1,480, since some individuals opted out, some email messages were undeliverable, and we incorrectly identified some authors. Of the 1,480 surveyed, 346 started the survey, but only 245 completed it sufficiently to be included in the analysis, for a response rate of 16.6%. Mean imputation, case-wise deletion, and pair-wise deletion were used during data analysis because some of the 245 surveys had missing data. The results were the same across all these methods. The final results reported in the following paragraphs are based on mean imputation.

5 | RESULTS

First, we assessed the psychometric properties of the scales for convergent and discriminant validity using factor analysis. During this process one question in each of the following three constructs—structural assurance, social factors, and continuance intention was eliminated given low factor loadings. All remaining items loaded more highly on their corresponding constructs than on the other constructs suggesting convergent validity within scales and discriminant validity across scales (Table 1). However, two items had loadings less than 0.7. Therefore, discriminant and convergent validity were

further assessed by examining the square root of the average variance shared (AVE) recommended by Fornell and Larcker (1981). The AVE is shown along the diagonals of the correlation matrix (Table 2). According to Fornell and Larcker (1981) values of 0.5 or higher indicate an acceptable level of convergent validity, while discriminant validity is indicated when the items of a construct share more variance internally than with other constructs in the model. The square root of the AVE of each construct was larger than its corresponding row and column correlations, indicating adequate discriminant validity (Table 2). All these multi-item measurement scales also showed high reliability, with Cronbach's alpha scores equal to or greater than 0.70 (Table 2).

Partial least squares was employed to empirically test the model using SmartPLS v4.08 to perform the analysis. We checked for multicollinearity and found that no variance inflation factor (VIF) approached the value of 3 indicating no evidence of multicollinearity (Cohen et al., 2003; Hair et al., 2010; Neter et al., 1996). Figure 1 shows the empirical examination of the entire model.

Unlike a covariance based structural equation model (CB-SEM), PLS-SEM does not seek to maximize a global criterion, resulting in the absence of global goodness-of-fit statistics (Hair et al., 2024). In addition, traditional CB-SEM fit measures also do not translate well to PLS-SEM because the interpretation of "fit" varies between CB-SEM and PLS-SEM (Hair et al., 2024; Henseler & Sarstedt, 2013). Within CB-SEM, "fit" is defined by how well the empirical data's covariance matrix is congruent with the theoretical model's implied covariance matrix. In PLS-SEM, "fit" reflects the congruence between actual and model-predicted values of the dependent variables (Hair et al., 2017, 2019; Henseler et al., 2014; Rigdon et al., 2017). Consequently, according to Hair et al. (2024) researchers should focus on a model's predictive strength rather than traditional CB-SEM fit metrics. In our study, the R-squared values for trust 64% ($p < 0.001$) and continuance intention 65% ($p < 0.001$) were both highly predictive, suggesting an appropriate fit.

Additionally, Smart PLS calculates the standardized root mean square residual (SRMR). SRMR has been proposed by Henseler et al. (2014) as a fitting criterion for PLS-SEM that avoids many of the limitations associated with the other CB-SEM fit measures. SRMR represents the mean absolute difference between observed correlations and those predicted by the model. Suitable model fit occurs with thresholds below 0.10 (liberal threshold) or 0.08 (more strict threshold) reflecting suitable model fit (Hair et al., 2024; Hu & Bentler, 1999). Our estimated model's SRMR of 0.078 suggests an appropriate fit.

Hypotheses 1, 4, and 5 were supported. Integrity (H1) ($\beta = 0.26$; $p < 0.01$), identification (H4) ($\beta = 0.25$; $p < 0.001$), and structural assurance (H5) ($\beta = 0.38$; $p < 0.001$) were positively related to data reuser trust. Hypotheses 2 and 3 were not supported. Benevolence (H2) ($\beta = 0.03$; $p > 0.05$) and transparency (H3) ($\beta = -0.01$; $p > 0.05$) were not positively related to data reuser trust. Hypotheses 6 and 7 were also supported. Both trust (H6) ($\beta = 0.34$; $p < 0.001$) and performance expectancy (H7) ($\beta = 0.50$; $p < 0.001$) were positively related to continuance intention. Social factors (H8) was not supported; it was not positively related to continuance intention ($\beta = 0.04$; $p > 0.05$). Finally, the three moderation hypotheses were not supported. Reuse reliance (H9) ($\beta = -0.09$; $p > 0.05$) and data scarcity (H10) ($\beta = -0.10$; $p > 0.05$) did not moderate the relationship between trust and continuance intention. Reuse experience (H11) ($\beta = 0.07$; $p > 0.05$) did not moderate the relationship between social factors and continuance intention.

6 | DISCUSSION

Motivated to investigate trust in a data repository from the perspective of its designated community of users, we created a model of trust formation and outcomes based on organizational, information systems, and archives and digital curation literature. Our findings demonstrate that data reusers bestow trust in data repositories when they believe it acts with integrity, demonstrates identification with scholarly practices, and exhibits structural assurance. Furthermore, trust, along with performance expectancy, is positively related to the trusting action, that is, continuance intention to use the data repository. Thus, data reuser trust in data repositories cannot be discounted and needs to be considered and investigated as part of the trust dynamic around the provision of digital information.

6.1 | Theoretical implications

Our study also extends theories of trust in the organizational and information systems literature. First, by examining the trustworthiness factors individually, we demonstrated that models originally developed in the for-profit sector including e-commerce can be applied in the non-profit environment and are therefore more generalizable. Our findings suggested data reusers' trust in a repository is based on concepts that speak to observable behavioral actions. Integrity is not simply stating a mission (i.e., preservation for long-term access to data), but acting in ways that demonstrate adherence to that mission. Identification encompasses learning the values and

TABLE 2 Means, standard deviations, reliabilities, and correlations matrix.

	Composite reliability	Mean (SD)	Benevolence	Continuance intention	Identification	Integrity	Performance expectancy	Social factors	Structural assurance	Transparency	Trust
Benevolence	0.77	5.44 (0.96)	(0.82)								
Continuance intention	0.80	6.22 (0.84)	0.49	(0.83)							
Identification	0.71	5.40 (0.94)	0.76	0.51	(0.77)						
Integrity	0.81	5.95 (0.92)	0.79	0.48	0.69	(0.85)					
Performance expectancy	0.85	6.13 (0.89)	0.54	0.75	0.66	0.56	(0.88)				
Social factors	0.72	5.26 (1.08)	0.46	0.44	0.54	0.48	0.48	(0.76)			
Structural assurance	0.73	5.64 (0.90)	0.62	0.58	0.65	0.65	0.58	0.57	(0.78)		
Transparency	0.82	5.59 (0.99)	0.79	0.53	0.73	0.73	0.57	0.44	0.64	(0.80)	
Trust	0.85	6.27 (0.78)	0.65	0.71	0.69	0.69	0.71	0.51	0.72	0.63	(0.88)

Note: Square root of the average variance extracted (AVE) is on the diagonal. Correlations above 0.13 are significant at the 0.05 level. $n = 245$.
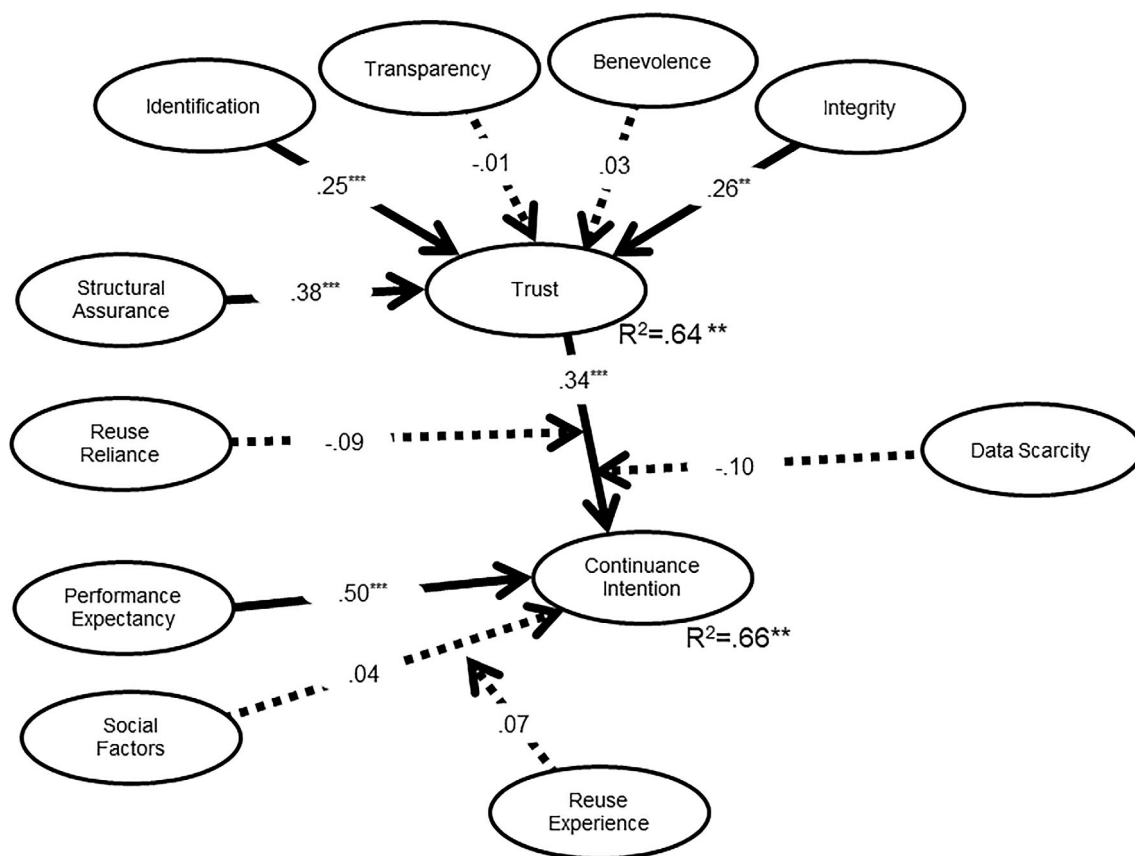
FIGURE 1   A model of data repository trust and continuance intention. **$p < 0.01$, ***$p < 0.001$.

practices of a designated community of users and enacting those values in the repository so that the community feels supported. Structural assurance which concerns the data repository's safety nets and other guarantees, such as third-party endorsements and reputation was also shown to have a positive relationship on data reuser trust. Since research has demonstrated that the curatorial work in repositories is invisible (Plantin, 2019; Thomer et al., 2022), structural assurance may be a proxy for transparency and/or benevolence.

Our study also makes a theoretical contribution to the archives and digital curation literature. Previous studies primarily focus on trust formation. We have attempted to go beyond trust formation to examine the trusting action. Furthermore, Prieto (2009) posits factors (although not always using these terms) that comprise trust in digital repositories: integrity, benevolence, identification, transparency, and structural assurance. We have taken Prieto's conceptualization, created a model and tested it to investigate trust formation and the ensuing trusting actions.

Finding both trust and performance expectancy were positively related to continuance intention lends additional evidence to prior information systems studies documenting the same phenomena (e.g., Akter et al., 2011; Bhattacherjee, 2001; He et al., 2009; Venkatesh

et al., 2003). Social factors (Hypothesis 8) was not supported. Prior research suggests people rely on their own perceptions more than those of their referent groups as they gain experience (e.g., Karahanna et al., 1999; Venkatesh et al., 2003; Venkatesh & Morris, 2000). Our finding may also be related to the fact that our respondents had an average 13.2 years of reuse experience.

None of the moderating factors changed the relationship between trust and continuance intention. All but one survey respondent experienced some level of reuse reliance. Yet, it had no influence on the relationship between trust in a data repository and their intention to continue using it. This may be because we asked about their reliance on "data produced by others." Asking specifically about the reliance on data housed in a repository might yield a different result. Data scarcity also did not change the relationship between trust and continuance intention. It may be that the degree to which survey respondents experienced data scarcity was not enough to limit their opportunities to engage with and build trust in data repositories. This aligns with previous findings by Borrás (2020) who investigated the underlying mechanisms for data reuse. She found that researchers reused data even when they had a less than an optimal amount. Lastly, reuse experience did not change the relationship

between social factors and continuance intention. All survey respondents had experience reusing data, which suggests reusing data over time may be less important than reusing data the first time. York (2022) found that reuse experience did not influence data reusers' need for more information about data. He posited that this may be partially explained by the fact that a substantial proportion of those who reuse data were mentored by or collaborated with others who had substantial knowledge of the data (York, 2022). The same may be true in our study with regard to repositories.

Finally, trust in repositories is just one element of trust in the digital information environment which encompasses the information purveyor (repository), staff, stakeholders, and the data itself. Due to the paucity of empirical research on trust in data repositories we examined this aspect of the environment. However, we see our findings as working towards a comprehensive model of trust which encompasses trust in data, trust in the curatorial staff who manage the data, stakeholders (data producers and reusers), and trust in the institutions that hold the data.

6.2 | Practical implications

If we are to support and sustain open access to data, repository staff attention must shift to the perspectives, practices, and needs of their designated communities of users. This shift would change the balance of trust factors as presented in OAIS, which largely focuses on the internal workings of a data repository, and consider what it looks like for a data repository to act in ways that communicate integrity, identification, and structural assurance such that data reusers want to bestow trust.

ICPSR demonstrates integrity and structural assurance by communicating the ways it fulfills its mission and adheres to internal and external standards, regulations, and safeguards. It exhibits identification as it works closely to understand and engage its designated communities by recognizing practices while encouraging high-quality data deposits and rigorous reuse that benefits society. Findings showed a high reward, but it takes a lot of effort and ICPSR is more well-resourced than others. It is difficult if not impossible for institutional repositories (IRs) to take on activities at this same level of effort, but they have found other ways to improve data practices throughout the lifecycle. In terms of increasing identification, IRs have partnered with libraries making it possible to offer data management workshops to understand and address the needs of their designated communities of users, schedule consultations, and work closely with

users at the point of data deposit. Moreover, these activities help data producers build their data management planning, organization, and curation knowledge to do things better the next time, such that future data deposits are of higher quality than first deposits. Prior research shows the same to be the case for data reusers who upon reflection change their own data creation, management, and curation practices in ways that yield higher quality data (Yakel et al., 2019). Moreover, data practices are learned through an apprenticeship process (e.g., Kriesberg et al., 2013), so it is likely that these producers and reusers will share their learning with students entering the field, who do it better their first time, thereby making it part of professional practice.

In the words of the OAIS (Consultative Committee for Space Data Systems, 2012), a repository's stated mission is to provide long-term access to data. However, recent work has been done to articulate the FAIR (findable, accessible, interoperable, reusable) and CARE (collective benefit, authority to control, responsibility, ethics) principles as well as operationalize and assess their implementation (e.g., Carroll et al., 2020, 2021, 2022; L'Hours et al., 2022; Research Data Alliance FAIR Data Maturity Model Working Group, 2020; Wilkinson et al., 2016). Criteria for selecting data repositories also has been introduced (National Institute of Health, 2020; Science Europe, 2021). Together, these activities suggest long-term access to data is not enough. One question is how our findings figure into these new FAIR and CARE principles and repository selection criteria. With more attention being brought to these principles and criteria, a repository must consider whether to incorporate them into its mission, and if so, how to demonstrate their impact in ways that meet the needs of certification bodies and resonate with designated communities of users. One way for repositories to keep pace with how these changes are impacting their designated communities' needs and practices is to turn to the research that has studied them and begin to incorporate the findings into practice (e.g., Johnston, 2020; Koesten et al., 2020; Trisovic et al., 2021).

6.3 | Limitations and future research

A model of trust in a data repository and the intention to continue using it was empirically tested. Although this is one of the first studies of its kind, it has several limitations. First, its designated community of users was narrowly focused on quantitative social scientists who reused data housed in ICPSR. One avenue for future research is testing the generalizability of these findings in

other disciplinary communities that reuse other types of data. Of particular interest might be designated communities where data reuse via data repositories is in the early stages. Given these different circumstances, different factors might come into play. For instance, social factors could succeed in igniting initial, if not continued, use of a data repository in its early years. Second, our examination of data reusers was in line with much of the technology acceptance literature that has examined information systems use from a consumer perspective (e.g., Gefen et al., 2003; Kimery & McCord, 2002), but data depositors were left out. Future research that examines whether, and if so how, the model of trust in a data repository tested in this study might change for data depositors is needed. Lastly, drawing from the OAIS (Consultative Committee for Space Data Systems) which describes a digital repository as an organization of people and systems, the data repository and the data provider were treated as one in our study. Likewise, some information systems studies show that users make a distinction between an information system and its provider and therefore the variables that precede and follow trust also vary (e.g., Söllner, 2020). Testing whether this holds true for a data repository is another area for future research that would expand the model of trust in a data repository to include performance, such as speed, reliability, availability, ease of use, and customization (Gefen et al., 2003; Lee & Turban, 2001; Söllner, 2020). Future research in these three areas would expand and contribute to the technology acceptance literature in general and the archives and data curation literature more specifically.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

ORCID

Elizabeth Yakel https://orcid.org/0000-0002-8792-6900
Ixchel M. Faniel https://orcid.org/0000-0001-7302-5936
Lionel P. Robert Jr https://orcid.org/0000-0002-1410-2601

REFERENCES

Adams, D. A., Nelson, R. R., & Todd, P. A. (1992). Perceived usefulness, ease of use, and usage of information technology: A replication. *MIS Quarterly*, *16*(2), 227–247. https://doi.org/10.2307/249577

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. https://doi.org/10.1002/asi.21442

Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pre-test assessment of their substantive validities. *Journal of Applied Psychology*, *76*(5), 732–740. https://doi.org/10.1037/0021-9010.76.5.732

Bak, G. (2016). Trusted by whom? TDRs, standards culture and the nature of trust. *Archival Science*, *16*(4), 373–402. https://doi.org/10.1007/s10502-015-9257-1

Barki, H., & Hartwick, J. (1994). Measuring user participation, user involvement, and user attitude. *MIS Quarterly*, *18*(1), 59–82. https://doi.org/10.2307/249610

Beagrie, N. (2013). Preservation, Trust and Continuing access for e-Journals (DPC Technology Watch Reports No. 4; pp. 1–49). Digital Preservation Coalition. http://dx.doi.org/10.7207/twr13-04

Bettivia, R. S. (2016). The power of imaginary users: Designated communities in the OAIS reference model. *Proceedings of the Association for Information Science and Technology*, *53*(1), 1–9. https://doi.org/10.1002/pra2.2016.14505301038

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*, *25*(3), 351–370. https://doi.org/10.2307/3250921

Borgman, C. L. (2015). *Big data, little data, no data: Scholarship in the networked world*. MIT Press. https://doi.org/10.7551/mitpress/9963.001.0001

Borrás, M. I. A. (2020). *A causal model to explain data reuse in science: A study in health disciplines* (Tesis doctoral). Universitat Politècnica de València. https://doi.org/10.4995/Thesis/10251/153164

Carlson, S., & Anderson, B. (2007). What are data? The many kinds of data and their implications for data re-use. *Journal of Computer-Mediated Communication*, *12*(2), 635–651. https://doi.org/10.1111/j.1083-6101.2007.00342.x

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE principles for indigenous data governance. *Data Science Journal*, *19*, 43. https://doi.org/10.5334/dsj-2020-043

Carroll, S. R., Garba, I., Plevel, R., Small-Rodriguez, D., Hiratsuka, V. Y., Hudson, M., & Garrison, N. A. (2022). Using

indigenous standards to implement the CARE principles: Setting expectations through tribal research codes. *Frontiers in Genetics*, 13, 823309. https://doi.org/10.3389/fgene.2022.823309

Carroll, S. R., Herczog, E., Hudson, M., Russell, K., & Stall, S. (2021). Operationalizing the CARE and FAIR principles for indigenous data futures. *Scientific Data*, 8(1), 108. https://doi.org/10.1038/s41597-021-00892-0

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates Publishers. Retrieved from https://psycnet.apa.org/record/2002-18109-000

Consultative Committee for Space Data Systems. (2011). Audit and certification of trustworthy digital repositories: Recommended practice (Standard ISO 16363:2012 (CCSDS 652.0-M-1)). Retrieved from https://www.iso.org/standard/56510.html

Consultative Committee for Space Data Systems. (2012). Space data and information transfer systems—Open archival information system (OAIS)—Reference model (Standard ISO 14721:2012 (CCSDS 650.0-M-2)). Retrieved from https://www.iso.org/standard/57284.html

Corrado, E. M. (2019). Repositories, trust, and the CoreTrustSeal. *Technical Services Quarterly*, 36(1), 61–72. https://doi.org/10.1080/07317131.2018.1532055

Cramer, A. N. (2022). Open data and the 21st-century naturalist. *Frontiers in Ecology and the Environment*, 20(10), 548–549. https://doi.org/10.1002/fee.2576

Cramer, H., Evers, V., Ramlal, S., van Someren, M., Rutledge, L., Stash, N., Aroyo, L., & Wielinga, B. (2008). The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction*, 18(5), 455–496. https://doi.org/10.1007/s11257-008-9051-3

Curty, R. (2015). Beyond "data thrifting": An investigation of factors influencing research data reuse in the social sciences. Dissertations—ALL. Retrieved from https://surface.syr.edu/etd/266

Curty, R. G., Crowston, K., Specht, A., Grant, B. W., & Dalton, E. D. (2017). Attitudes and norms affecting scientists' data reuse. *PLoS One*, 12(12), e0189288. https://doi.org/10.1371/journal.pone.0189288

Daniels, M., Faniel, I., Fear, K., & Yakel, E. (2012). Managing fixity and fluidity in data repositories. In *Proceedings of the 2012 iConference* (pp. 279–286). ACM. https://doi.org/10.1145/2132176.2132212

Darch, P. T. (2018). Limits to the pursuit of reproducibility: Emergent data-scarce domains of science. In *Transforming digital worlds: 13th international conference, iConference 2018 Sheffield, UK, March 25–28, 2018 Proceedings* (pp. 164–174). Springer.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. https://doi.org/10.2307/249008

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the Workplace1. *Journal of Applied Social Psychology*, 22(14), 1111–1132. https://doi.org/10.1111/j.1559-1816.1992.tb00945.x

Edelman. (2022). 2022 Edelman Trust Barometer. Retrieved from https://www.edelman.com/trust/2022-trust-barometer

Edwards, P. N. (2010). *A vast machine: Computer models, climate data, and the politics of global warming*. MIT Press.

Faniel, I. M., & Jacobsen, T. E. (2010). Reusing scientific data: How earthquake engineering researchers assess the reusability of Colleagues' data. *Computer Supported Cooperative Work (CSCW)*, 19(3), 355–375. https://doi.org/10.1007/s10606-010-9117-8

Faniel, I. M., Kansa, E., Whitcher Kansa, S., Barrera-Gomez, J., & Yakel, E. (2013). The challenges of digging data: A study of context in archaeological data reuse. In *Proceedings of the 13th ACM/IEEE-CS joint conference on digital libraries* (pp. 295–304). ACM. https://doi.org/10.1145/2467696.2467712

Faniel, I. M., Kriesberg, A., & Yakel, E. (2012). Data reuse and sensemaking among novice social scientists: Data reuse and sensemaking among novice social scientists. *Proceedings of the American Society for Information Science and Technology*, 49(1), 1–10. https://doi.org/10.1002/meet.14504901068

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, 18(3), 382–388. https://doi.org/10.2307/3150980

Frank, R. D., Yakel, E., & Faniel, I. M. (2015). Destruction/reconstruction: Preservation of archaeological and zoological research data. *Archival Science*, 15(2), 141–167. https://doi.org/10.1007/s10502-014-9238-9

Garrett, J., & Waters, D. (1996). *Preserving digital information: Report of the task force on archiving of digital information*. Commission on Preservation and Access and the Research Libraries Group.

Gefen, D., Karahanna, E., & Straub, D. (2003). Trust and TAM in online shopping: An integrated model. *Management Information Systems Quarterly*, 27(1), 51.

Gefen, D., & Straub, D. W. (2004). Consumer trust in B2C e-commerce and the importance of social presence: Experiments in e-products and e-services. *Omega*, 32(6), 407–424. https://doi.org/10.1016/j.omega.2004.01.006

Grimmelikhuijsen, S. (2009). Do transparent government agencies strengthen trust? *Information Polity*, 14(3), 173–186. https://doi.org/10.3233/IP-2009-0175

Groves, R. M., Fowler, F. J., Couper, M., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology*. Wiley.

Hair, J., Hollingsworth, C. L., Randolph, A. B., & Chong, A. Y. L. (2017). An updated and expanded assessment of PLS-SEM in information systems research. *Industrial Management & Data Systems*, 117(3), 442–458. https://doi.org/10.1108/IMDS-04-2016-0130

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. (2010). *Multivariate data analysis* (5th ed.). Prentice Hall.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hair, J. F., Sharma, P. N., Sarstedt, M., Ringle, C. M., & Liengaard, B. D. (2024). The shortcomings of equal weights estimation and the composite equivalence index in PLS-SEM. *European Journal of Marketing*, 58(13), 30–55. https://doi.org/10.1108/EJM-04-2023-0307

Hallikainen, H., & Laukkanen, T. (2021). Trustworthiness in e-commerce: A replication study of competing measures. *Journal of Business Research*, 126, 644–653. https://doi.org/10.1016/j.jbusres.2019.12.033

He, W., Fang, Y., & Wei, K.-K. (2009). The role of trust in promoting organizational knowledge seeking using knowledge

management systems: An empirical investigation. *Journal of the American Society for Information Science and Technology*, *60*(3), 526–537. https://doi.org/10.1002/asi.21006

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen, D. J., Hair, J. F., Hult, G. T. M., & Calantone, R. J. (2014). Common beliefs and reality about PLS: Comments on Rönkkö and Evermann (2013). *Organizational Research Methods*, *17*(2), 182–209. https://doi.org/10.1177/1094428114526928

Henseler, J., & Sarstedt, M. (2013). Goodness-of-fit indices for partial least squares path modeling. *Computational Statistics*, *28*(2), 565–580. https://doi.org/10.1007/s00180-012-0317-1

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jantz, R., & Giarlo, M. (2007). Digital archiving and preservation: Technologies and processes for a trusted repository. *Journal of Archival Organization*, *4*(1–2), 193–213. https://doi.org/10.1300/J201v04n01_10

Johnston, L. R. (2020, February 13). Librarians can help leverage research data in new ways. *OCLC Next*. Retrieved from https://blog.oclc.org/next/data-curators-network/

Karahanna, E., Straub, D. W., & Chervany, N. L. (1999). Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. *MIS Quarterly*, *23*(2), 183–213. https://doi.org/10.2307/249751

Kim, B., & Han, I. (2009). The role of trust belief and its antecedents in a community-driven knowledge environment. *Journal of the American Society for Information Science and Technology*, *60*(5), 1012–1026. https://doi.org/10.1002/asi.21041

Kimery, K., & McCord, M. (2002). Third party assurances: Mapping the road to trust in eRetailing. *Journal of Information Technology Theory and Application (JITTA)*, *4*(2), 63–82.

Koesten, L., Vougiouklis, P., Simperl, E., & Groth, P. (2020). Dataset reuse: Toward translating principles to practice. *Patterns*, *1*(8), 100136. https://doi.org/10.1016/j.patter.2020.100136

Komiak, S., & Benbasat, I. (2008). A two-process view of trust and distrust building in recommendation agents: A process-tracing study. *Journal of the Association for Information Systems*, *9*(12), 727–747. https://doi.org/10.17705/1jais.00180

Komiak, S. Y. X., & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly*, *30*(4), 941–960. https://doi.org/10.2307/25148760

Kriesberg, A., Frank, R. D., Faniel, I. M., & Yakel, E. (2013). The role of data reuse in the apprenticeship process: The role of data reuse in the apprenticeship process. *Proceedings of the American Society for Information Science and Technology*, *50*(1), 1–10. https://doi.org/10.1002/meet.14505001051

Lee, M. K. O., & Turban, E. (2001). A trust model for consumer internet shopping. *International Journal of Electronic Commerce*, *6*(1), 75–91. https://doi.org/10.1080/10864415.2001.11044227

Lewicki, R. J., & Bunker, B. B. (1995). Trust in relationships: A model of trust development and decline. In M. Deutsch, B. B. Bunker, & J. Z. Rubin (Eds.), *Conflict, cooperation, and justice: essays inspired by the work of Morton Deutsch*. Jossey-Bass Publishers.

L'Hours, H., Verburg, M., De Vries, J., Cepinskas, L., Von Stein, I., Huber, R., Davidson, J., Herterich, P., & Mathers, B. (2022). Report on a maturity model towards FAIR data in FAIR repositories (D4.6). https://doi.org/10.5281/ZENODO.6090388

Li, X., Hess, T. J., & Valacich, J. S. (2008). Why do we trust new technology? A study of initial trust formation with organizational information systems. *Journal of Strategic Information Systems*, *17*(1), 39–71. https://doi.org/10.1016/j.jsis.2008.01.001

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

McKnight, H. D., Choudhury, V., & Kacmar, C. (2002). The impact of initial consumer trust on intentions to transact with a web site: A trust building model. *Journal of Strategic Information Systems*, *11*(3), 297–323. https://doi.org/10.1016/S0963-8687(02)00020-3

Meeßen, S. M., Thielsch, M. T., & Hertel, G. (2020). Trust in management information systems (MIS). *Zeitschrift Für Arbeits-Und Organisationspsychologie A&O*, *64*(1), 6–16. https://doi.org/10.1026/0932-4089/a000306

National Institutes of Health. (2020). Supplemental information to the NIH policy for data management and sharing: Selecting a repository for data resulting from NIH-supported research. Retrieved from https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-016.html

National Institutes of Health. (2023, January 25). Final NIH policy for data management and sharing. Retrieved from https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html

Nature, S. (n.d.). Springer Nature discovery services. Retrieved from https://www.springernature.com/gp/librarians/tools-services/implement/discovery/discovery-services

Neter, J., Kutner, M., & Wasserman, W. (1996). *Applied linear statistical models* (4th ed.). Irwin.

Oliver, G., Chawner, B., & Liu, H. (2011). Implementing digital archives: Issues of trust. *Archival Science*, *11*(3), 311–327. https://doi.org/10.1007/s10502-011-9167-9

Pew Research Center. (2020). *Americans' views of government: Low trust, but some positive performance ratings*. Pew Research Center-U.S. Politics & Policy. Retrieved from https://www.pewresearch.org/politics/2020/09/14/americans-views-of-government-low-trust-but-some-positive-performance-ratings/

Pirson, M., & Malhotra, D. (2011). Foundations of organizational trust: What matters to different stakeholders? *Organization Science*, *22*(4), 1087–1104. https://doi.org/10.1287/orsc.1100.0581

Plantin, J.-C. (2019). Data cleaners for pristine datasets: Visibility and invisibility of data processors in social science. *Science, Technology, & Human Values*, *44*(1), 52–73. https://doi.org/10.1177/0162243918781268

Price, D. M., & Smith, J. J. (2011). The trust continuum in the information age: A Canadian perspective. *Archival Science*, *11*(3–4), 253–276. https://doi.org/10.1007/s10502-011-9148-z

Prieto, A. G. (2009). From conceptual to perceptual reality: Trust in digital repositories. *Library Review*, *58*(8), 593–606. https://doi.org/10.1108/00242530910987082

Research Data Alliance FAIR Data Maturity Model Working Group. (2020). *FAIR data maturity model: Specification and guidelines*. Research Data Alliance. https://doi.org/10.15497/RDA00050

Research Libraries Group-OCLC (RLG/OCLC) Working Group on Digital Archive Attributes. (2002). *Trusted digital repositories: Attributes and responsibilities: An RLG-OCLC report* (pp. 1–70). Research Libraries Group. https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf

Ribes, D., & Bowker, G. C. (2009). Between meaning and machine: Learning to represent the knowledge of communities. *Information and Organization*, *19*(4), 199–217. https://doi.org/10.1016/j.infoandorg.2009.04.001

Rigdon, E. E., Sarstedt, M., & Ringle, C. M. (2017). On comparing results from CB-SEM and PLS-SEM: Five perspectives and five recommendations. *Marketing ZFP*, *39*(3), 4–16. https://doi.org/10.15358/0344-1369-2017-3-4

Ross, S., & McHugh, A. (2006). Preservation pressure points: Evaluating diverse evidence for risk management. Retrieved from https://ecommons.cornell.edu/handle/1813/3674

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Introduction to special topic forum: Not so different after all: A cross-discipline view of trust. *The Academy of Management Review*, *23*(3), 393–404.

Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods*, *15*(1), 85–109. https://doi.org/10.1177/1525822X02239569

Saad, L. (2023, July 6). Historically Low Faith in U.S. Institutions Continues. Gallup.com. https://news.gallup.com/poll/508169/historically-low-faith-institutions-continues.aspx

Schäfer, A., Pampel, H., Pfeiffenberger, H., Dallmeier-Tiessen, S., Tissari, S., Darby, R., Giaretta, K., Giaretta, D., Gitmans, K., Helin, H., Lambert, S., Mele, S., Reilly, S., Ruiz, S., Sandberg, M., Schallier, W., Schrimpf, S., Smit, E., Wilkinson, M., & Wilson, M. (2011). Baseline report on drivers and barriers in data sharing. Opportunities for data exchange (ODE). Retrieved from http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2011/11/ODE-WP3-DEL-0002-1_0_public_final.pdf

Schultz, M., & Gore, E. (2010). The importance of trust in distributed digital preservation: A case study from the metaarchive cooperative. In *Proceedings of the 7th international conference on preservation of digital objects* (pp. 105–112). Österreichische Computer Gesellschaft. Retrieved from http://ifs.tuwien.ac.at/dp/ipres2010/papers/schultz-39.pdf

Science Europe. (2021). *Practical guide to the international alignment of research data management—Extended edition*. Science Europe. https://doi.org/10.5281/ZENODO.4915861

Serva, M. A., Fuller, M. A., & Mayer, R. C. (2005). The reciprocal nature of trust: A longitudinal study of interacting teams. *Journal of Organizational Behavior*, *26*(6), 625–648. https://doi.org/10.1002/job.331

Shankar, K., Eschenfelder, K. R., & Downey, G. (2016). Studying the history of social science data archives as knowledge infrastructure. *Science & Technology Studies*, *29*, 62–73.

Sitkin, S. B., & Roth, N. L. (1993). Explaining the limited effectiveness of legalistic "remedies" for trust/distrust. *Organization Science*, *4*(3), 367–392. https://doi.org/10.1287/orsc.4.3.367

Söllner, M. (2020). Same same but different? A two-foci perspective on trust in information systems. In *Hawaii international conference on system sciences*. Shidler College of Business. https://doi.org/10.24251/HICSS.2020.630

Speck, J. G. (2010). Protecting public trust: An archival wake-up call. *Journal of Archival Organization*, *8*(1), 31–53. https://doi.org/10.1080/15332748.2010.483389

Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information Systems*, *13*, 1. https://doi.org/10.17705/1CAIS.01324

Taylor, S., & Todd, P. A. (1995). Understanding information technology usage: A test of competing models. *Information Systems Research*, *6*(2), 144–176. https://doi.org/10.1287/isre.6.2.144

Thomer, A. K., Akmon, D., York, J. J., Tyler, A. R. B., Polasek, F., Lafia, S., Hemphill, L., & Yakel, E. (2022). The craft and coordination of data curation: Complicating workflow views of data science. *Proceedings of the ACM on Human-Computer Interaction*, *6*, 1–29. https://doi.org/10.1145/3555139

Thompson, R. L., Higgins, C. A., & Howell, J. M. (1991). Personal computing: Toward a conceptual model of utilization. *MIS Quarterly*, *15*(1), 125–143. https://doi.org/10.2307/249443

Triandis, H. C. (1980). Values, attitudes, and interpersonal behavior. *Nebraska Symposium on Motivation. Nebraska Symposium on Motivation*, *27*, 195–259.

Trisovic, A., Mika, K., Boyd, C., Feger, S., & Crosas, M. (2021). Repository approaches to improving the quality of shared data and code. *Data*, *6*(2), 15. https://doi.org/10.3390/data6020015

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MIS Quarterly*, *24*(1), 115–139. https://doi.org/10.2307/3250981

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Wang, W., & Benbasat, I. (2008). Attributions of trust in decision support technologies: A study of recommendation agents for E-commerce. *Journal of Management Information Systems*, *24*(4), 249–273. https://doi.org/10.2753/MIS0742-1222240410

White House Office of Science and Technology Policy. (2023). Report to the U.S. congress on financing mechanisms for open access publishing of federally funded research. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2023/11/Open-Access-Publishing-of-Scientific-Research.pdf

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Santos, L. B. S., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018. https://doi.org/10.1038/sdata.2016.18

Xie, H., Prybutok, G., Peng, X., & Prybutok, V. (2020). Determinants of trust in health information technology: An empirical investigation in the context of an online clinic appointment system. *International Journal of Human–Computer Interaction*, *36*(12), 1095–1109. https://doi.org/10.1080/10447318.2020.1712061

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

Yakel, E., Faniel, I. M., & Maiorana, Z. J. (2019). Virtuous and vicious circles in the data life-cycle. Retrieved from https://informationr.net/ir/24-2/paper821.html

Yoon, A. (2013). Defining what matters when preserving web-based personal digital collections: Listening to bloggers. *International Journal of Digital Curation*, *8*(1), 173–192. https://doi.org/10.2218/ijdc.v8i1.240

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

Yoon, A. (2015). *Data reuse and users' trust judgments: Toward trusted data curation-ProQuest*. The University of North Carolina at Chapel Hill. Retrieved from https://www.proquest.com/docview/1718199510?pq-origsite=gscholar&fromopenview=true

Yoon, A. (2016). Visible evidence of invisible quality dimensions and the role of data management. https://doi.org/10.9776/16123

Yoon, A., & Lee, Y. Y. (2019). Factors of trust in data reuse. *Online Information Review*, *43*(7), 1245–1262. https://doi.org/10.1108/OIR-01-2019-0014

York, J. (2022). *Seeking equilibrium in data reuse: A study of knowledge satisficing* (Thesis). https://doi.org/10.7302/6170