Data, Decisions, and Doubt: Affording Trust Within Environmental Data Science

Lauren Victoria Mary Thornton, BA (Hons), MA, MSc

School of Computing and Communications

Lancaster University

This dissertation is submitted for the degree of

Doctor of Philosophy

December, 2024

For my sunshine boy.

Data, Decisions, and Doubt: Affording Trust Within Environmental Data Science

Lauren Victoria Mary Thornton, BA (Hons), MA, MSc.

School of Computing and Communications, Lancaster University

A thesis submitted for the degree of *Doctor of Philosophy.* December, 2024.

Abstract

Trust is recognised as a valuable element of socio-technical systems, facilitating numerous important benefits. Yet despite extensive study within and across multiple disciplines, achieving well-placed trust remains a perpetual challenge. This research explores the perception of trust within environmental data science: transdisciplinary, collaborative scientific research responding to fundamentally complex, global problems through data science methodologies and techniques. In this context, the presence of trust can reduce uncertainty and friction, consequently facilitating collaboration, generating robust scientific results, and supporting the reuse of data and models. To begin, this research undertakes a theoretical analysis of trust, deconstructing the typical foci of research – models of trust – arguing instead for a different modality: trust affordances. Trust affordances account for trust within system design by attending to multiple stakeholders arriving at trust from varying angles. As a concept, trust affordances are characteristics of the technology by virtue of itself or of features designed into the technology to promote trust by providing access to evidence of (dis)trustworthiness specific to a user, a technology, and their context.

This work concentrates on the perceptions of stakeholders across the data science pipeline within two cases. In the first, six interviews and a workshop attended by seventeen participants are used to consider the perceptions of trust and trustworthiness by data curators within a Research Council. In the second case, thirteen interviews are undertaken and a survey of ten respondents are used

to generate the views of data producers and consumers related to a Centre of Excellence. The empirical data is analysed twice to respectively produce general themes related to trust and instances of trust affordances. This research contends that whilst there is no universal answer or one-size-fits-all approach to trust, system designers can actively design for trust by embracing nuance and complexity with a creative, thorough, and thoughtful approach.

Acknowledgements

This research would not have been possible without the funding provided by the Data Science Institute, the School of Computing and Communications, and the EPSRC (grant number: EP/R512564/1). Equally, this research would not have been possible without the participants who agreed to take part in this work, thank you for your time and the insights provided. A thank you is also needed for my examiners, Corina Sas and Steve Marsh, for a providing a positive viva experience and contributing their insights. Thanks also to Jan Bebbington, for providing me with the opportunity to carry on with this research in the Pentland Centre for Sustainability in Business and for giving me the time and space to finish this work.

This research would not – in any shape or form – have been possible without the efforts of my doctoral supervisors, Bran Knowles and Gordon Blair. Thanks for always cheering me on and providing me with constructive criticism to improve my work. As this is a public record, you both know that I *won't* be saying anything sentimental or exposing our inside jokes, but I'm sure you already know how immensely appreciative and grateful I am for not only giving me this opportunity but also encouraging me to follow my ideas and to pursue post-doc research.

Thanks also to the friends, family, colleagues, and acquaintances that were there before and during this journey or who have joined along the way. If I had known what the Fates had in store for me at the start of this degree, I probably wouldn't have started, but you all either physically, mentally, or emotionally patched me up and made sure that I could make it to the end.

The acknowledgements wouldn't be complete without mentioning Ted (Teddy, Theodore, Beddy, Bebbs, Teebs, all of the bears (Barnaby, Paddington, Rupert, Bartholomew, and *the* Teddy Edward) and all of your other nicknames). Thank you for being the leader of our pack and for being my companion and co-conspirator, for bringing peace and purpose, love and laughter, and simply for being there by my side, with your warmth and light even in the darkest times.

Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and belief, original and my own work. The material has not been submitted, either in whole or in part, for a degree at this, or any other university. This thesis does not exceed the maximum permitted word length of 80,000 words including appendices and footnotes, but excluding the bibliography.

Lauren Victoria Mary Thornton

Publications

Contributing Publications

Several peer-reviewed publications, shown below, have been created directly from the research which this thesis describes.

Thornton, L., Neumann, V., Blair, G., Davies, N., and Watkins, J. (2019). "Trusted Brokers?: Identifying the Challenges Facing Data Centres". *Data for Policy*. London, UK, pp. 1-6. **This paper contributes to chapter 5.**

Thornton, L., Knowles, B., and Blair, G. (2021). "Fifty Shades of Grey: In Praise of a Nuanced Approach Towards Trustworthy Design". In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*. Virtual Event, Canada: Association for Computing Machinery, pp. 64–76. **This paper contributes to chapter 3.**

Thornton, L., Knowles, B., and Blair, G. (2022). "The Alchemy of Trust: The Creative Act of Designing Trustworthy Socio-Technical Systems". In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. Seoul, Republic of Korea: Association for Computing Machinery, pp. 1387–1398. **This paper contributes to chapter 2.**

Additional Publications

The following peer-reviewed publications have been generated whilst conducting this doctoral research.

Blair, G., Bassett, R., [..] **Thornton, L.** et al. (2021). "The Role of Digital Technologies in Responding to the Grand Challenges of the Natural Environment: The Windermere Accord", *Patterns* 2(1), pp. 1-8.

Sturdee, M., **Thornton, L.**, Wimalasiri, B., and Patil, S. (2021). "A Visual Exploration of Cybersecurity Concepts". In: *Proceedings of the 13th Conference on Creativity and Cognition (C & C '21)*. Virtual Event, Italy: Association for Computing Machinery, pp. 1-10.

Glossary

Affordance An affordance is a perceived property of an object which guides the user to operate or use it, e.g., a scrollbar on a computer screen is not real, but the shading and arrows connote that one can scroll down or up.

Environmental Data Science A blend of both environmental science (and the many sub-disciplines within it) with data science, which includes the application of statistical and computational methods to environmental science problems.

Model Refers to the representation of real-world processes, such as a conceptual model of trust. In environmental science, 'model' is usually used to refer to either stochastic modelling, e.g., a simulation of climate change processes or statistical modelling, e.g., a time series of crop yields collected within a field.

Transdisciplinary Research This style of research involves a combination of academic researchers across multiple disciplines and importantly, includes engagement with stakeholders beyond academia throughout the whole research process.

Trust A positive expectation wherein the trustor – in a position of risk and vulnerability – believes that the trustee will uphold and act in the interest of the trustor.

Trustee The person, group, institution, or object in which trust is placed.

Trustor The person placing trust in a trustee or trustees.

Trustworthiness The qualities a person or object is perceived to possess which makes them able to be trusted in the eyes of a trustor.

Trust Affordance Provides the opportunity for a trustor to place trust in the trustworthy. Trust affordances exist a higher level than features, but can be used to guide system designers in developing technology that is trustworthy-by-design.

Virtual Lab Virtual research environments that enable communication and collaboration between users. They provide a flexible infrastructure and access to cloud computing with support for integration of multiple models and systems, making them customisable to addressing different grand challenges.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Trust "brings us all sorts of good things" (Uslaner, 2000, p.569) enabling small and large acts that would otherwise be impossible or unavailable without trust (Botsman, 2017; Luhmann, 2017). Trust aids in forming positive expectations, making decisions, and enabling a willingness to act in uncertain, complex, or risky scenarios (Barber, 1983; Giddens, 1990; Lewis and Weigert, 1985; Luhmann, 2017; Mayer, Davis, and Schoorman, 1995; Rousseau et al., 1998). The prevalence of trust within multiple disciplines from economics to social science is indicative of these 'good things' that trust is thought to bring about (Castelfranchi and Falcone, 2016; Söllner et al., 2012; Van House, 2002; Yoon, 2017). This research focuses on trust in environmental data science: holistic and collaborative transdisciplinary research to address complex and time-sensitive global environmental challenges (Blair, Leeson, et al., 2019; Boersema, 2009; Easterbrook, 2010; Holzer et al., 2019; Mattor et al., 2014).

Environmental data science is comprised *precisely* of these 'small and large acts' for which trust is so beneficial. Results must be robustly trustworthy to stand up against competing values and politics at regional, national, and global scales (Berners-Lee, 2019; Edwards, 2010; Lucas, Leith, and Davison, 2015; O'Neill, 2002; Oreskes, 2019). However there are also issues of complexity, uncertainty, and risk at a micro-scale level, including data uncertainty and veracity (Blair, Leeson, et al., 2019;

Hays, 2000). Data is used, handled, and selected for different purposes. Data may be combined to form datasets which feed into scientific models or used to 'ground-truth' new empirical data (Pasquetto, Borgman, and Wofford, 2019; Pescott, Humphrey, and Walker, 2018; Vaughan and Wallach, 2021; Wallis, Borgman, et al., 2007). Thus trust is needed at various stages in the data pipeline including the selection, reuse and sharing of data, models and knowledge (Marsh, Atele-Williams, et al., 2020; O'Neill, 2018; Penner and Klahr, 1996; Van House, Butler, and Schiff, 1998; Van House, 2002). Trust also extends to the social, facilitating and sustaining collaborative relationships (Dwyer, Basu, and Marsh, 2013; Marsh, Atele-Williams, et al., 2020; McKnight, Choudhury, and Kacmar, 2002; Meyerson, Weick, and Kramer, 1996). This thesis will focus upon two main challenges to fostering trust in scientific collaboration: distributed virtual teamwork (Birnholtz and Bietz, 2003; Latour and Woolgar, 1979; Olson, Teasley, et al., 2002; Rolin, 2020) and working with unfamiliar data, models and people (Cragin and Shankar, 2006; Sundqvist, 2011; Van House, Butler, and Schiff, 1998; Van House, 2002).

Research on trust within environmental data science is both timely and important, as demonstrated by the inclusion of 'facilitating confidence and trust in digital research' as one of the seven themes for development in the National Environment Research Council (NERC) Digital Strategy for 2021-2030 (Natural Environment Research Council, 2022). Within this theme, NERC states that "for data and digital technologies to be used effectively by the environmental science community they *must be trustworthy* and reliable, traceable, accessible and usable" with their ambition being to "*facilitate confidence and trust in digital research*, and to ensure that NERC's data assets can be used and re-used with confidence by colleagues from *across and beyond the environmental science community*, by other disciplines and across sectors" (Natural Environment Research Council, 2022, p.14, emphasis added). Given this, trust is undoubtedly important, but how may we 'facilitate trust' or 'demonstrate trustworthiness' in data and technology for these varied stakeholders?

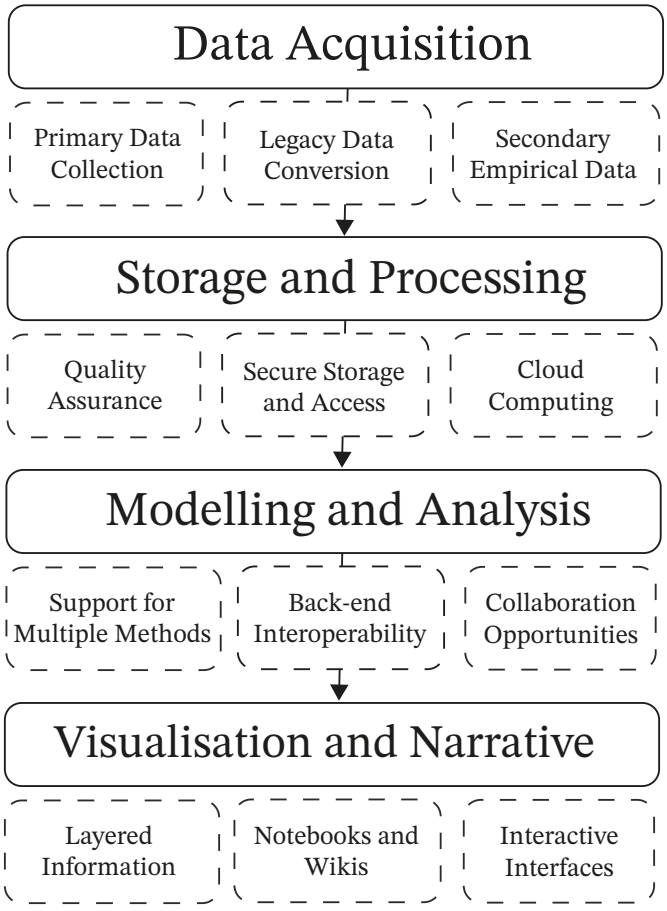


Figure 1.1: The design of a Virtual Lab can be configured in multiple ways. This illustration shows the data science pipeline (solid lines) which examples of features within each (dotted lines).

The design of technological infrastructure represents one avenue in addressing these questions, particularly by fostering communication, collaboration, and knowledge transfer (Birnholtz and Bietz, 2003; Faniel and Zimmerman, 2011; Jirotka et al., 2005). This research builds upon investigations into the role of technical systems to support open, honest, collaborative, and reproducible science — of which *trust* is recognised as playing a key part (Birnholtz and Bietz, 2003; Blair, Bassett, et al., 2021; Faniel and Zimmerman, 2011; Olson, Teasley, et al., 2002). This research focuses specifically on *Virtual Labs*, depicted in figure 1.1, as a tool that

can promote trust (Blair, Bassett, et al., 2021; Hollaway et al., 2020).[1] Virtual Labs are cloud-based platforms enabling users to combine data and modelling software through a Web browser.[2] These technologies can act as a bridge between communities (Finholt and Olson, 1997; Jeffery, Candela, and Glaves, 2020) including those working outside of scientific academic research (Blair, Leeson, et al., 2019; Elkhatib, Blair, and Surajbali, 2013; Lucas, Leith, and Davison, 2015; Mattor et al., 2014; Zuiderwijk, 2017). Virtual Labs have potential for addressing many of these challenges, including facilitating trust (European Commission, 2015), yet little is currently known regarding the design of such systems in order to effectively foster trust.

1.1 Research Questions

This research explores the role of Virtual Labs in fostering trust and trustworthiness within environmental data science. The overall aim of the research is to **inform the design of technology to support well-placed, appropriate trust for multiple stakeholders whilst embracing the nuances and complexities of trust within environmental data science**. There are several key considerations to note with respect to this aim, which form the research questions this work will address:

> **How do stakeholders determine trustworthiness in order to place or refuse trust within environmental data science?** The common assertion that trust is important and beneficial is based upon the assumption that trust is well-placed. However, the focus is better

[1] Also referred to as virtual research environments (Candela, Castelli, and Pagano, 2013; De Roure, Goble, Bhagat, et al., 2008), digital libraries (Schiff, Van House, and Butler, 1997; Van House, Bishop, and Buttenfield, 2003), and collaboratories (Birnholtz and Bietz, 2003; Chin and Lansing, 2004; Finholt and Olson, 1997; Zimmerman, Bos, et al., 2009).

[2] Examples include `https://bccvl.org.au`, `https://cordis.europa.eu/project/id/676247` and `https://datalab.datalabs.ceh.ac.uk` accessed July 2023.

understood as *trust in the trustworthy and refusal of trust in the untrustworthy*, i.e., a correct placement of trust (O'Hara, 2019; O'Neill, 2018; Riegelsberger, Sasse, and McCarthy, 2005). Trust is beneficial in certain conditions, but not *all* conditions. This is often taken for granted, distrust may be appropriate when something or someone is untrustworthy. Ultimately, one would want to empower users to make choices by aligning trustworthiness cues with their perceptions (Dwyer, Basu, and Marsh, 2013; Dwyer and Marsh, 2016; O'Hara, 2019; Marsh, Atele-Williams, et al., 2020; Riegelsberger, Sasse, and McCarthy, 2003b), bringing perceptions of trustworthiness in line with trustworthy trustees. Precisely how this is achieved, given that there is no agreement on what 'trusted' means in itself (Button, 2006; Knowles, 2016; Misztal, 2020) *and* more importantly, what trust means to those served within the digital data curation space (Jantz and Giarlo, 2007; Prieto, 2009), is thus the first key consideration of this research.

What are the salient types of trust within environmental data science and how do they differ? Consideration is needed of the trustees (that in which trust may be placed) and the trustors (those placing trust). There are multiple trustors trusting: data, people, and systems all of which are interrelated and interlinked (Jantz and Giarlo, 2007; Knowles, Rouncefield, et al., 2015; Knowles, 2016; Marsh, Atele-Williams, et al., 2020). Trust in any one trustee *necessarily implies* trust in other interconnected elements either as a mediator or a facilitator of trust, e.g., trust between people mediated by technology, and between people and technology (Jirotka et al., 2005; Sas and Khairuddin, 2017; Söllner et al., 2012). Though these interconnections are not always made explicit. A system must be deemed trustworthy in order for users to view the tools and information within the system to be perceived as trustworthy (Van House, 2003). Relatedly, there

are multiple stakeholders within environmental data science traversing different times, places, and disciplines (Bates, Lin, and Goodale, 2016; Carlson and Anderson, 2007; Faniel and Zimmerman, 2011; Zimmerman, 2008a). This raises specific challenges when designing for trust that may not be present in other contexts, e.g., to fit the needs of multiple communities (Barker et al., 2019; Cragin and Shankar, 2006) for present, known users and future, unknown users (Bates, Lin, and Goodale, 2016; Carlson and Anderson, 2007; Faniel and Zimmerman, 2011; Zimmerman, 2008a). There is a need to *balance* the different perspectives of those placing their trust (Sexton et al., 2005) and consider whether there are any tensions between multiple trusts that potentially compete (Dwyer and Marsh, 2016; Grudin, 1994a; Knowles, Rouncefield, et al., 2015; Pirson and Malhotra, 2010; Sexton et al., 2005; Van House, Butler, and Schiff, 1998). These trusts are different, engaging different forms of evidence for different stakeholders. Appealing equally to these trusts is complex, and currently how these trusts function within environmental data science is unclear.[3] This research seeks to contribute to the design of Virtual Labs such that benefits are created for multiple trustors (cf. Grudin, 1994b) and the multiple trustees in which they place trust by explicitly considering these interrelations and conflicts.

How can considerations of trust be incorporated into system design? Trust as a concept itself appears to defy a *truly universal* definition (Blomqvist, 1997; Dwyer and Marsh, 2016; Gambetta, 1988; Misztal, 2020): it is polysemous, containing varied meanings (Lahno, 2020; Riegelsberger, Sasse, and McCarthy, 2003b; Sundqvist, 2011); has been studied variously across disciplines (Misztal, 2020; McKnight, Choudhury, and Kacmar, 2002; Shapiro, 1987); and, is in itself dynamic,

[3]The use of the term 'trusts' is included specifically to highlight multiple forms of trust within a context.

changing based upon the experiences and contexts of the trustor (Lewicki and Bunker, 1996; McKnight, Carter, et al., 2011; Möllering, 2005; Schilke and Cook, 2013). Considering trust from the outset and throughout when designing a system (Knowles, Harding, et al., 2014) is challenging: developing a system of *any* type isn't obvious when it comes to trust. There is *no blueprint for design* that can be applied to each and every technology (Ashoori and Weisz, 2019; Easterbrook, 2014; Marsh, Atele-Williams, et al., 2020), and there *are limited design principles, frameworks or approaches* to utilise (Gefen, Benbasat, and Pavlou, 2008; Jacovi et al., 2021; Knowles, Rouncefield, et al., 2015; Lin et al., 2020). There is a gap in knowledge about how to design a system for well-placed trust that takes into account these complexities and nuances. This research is inspired by the notion of 'trustworthy by design' (Knowles, Harding, et al., 2014), that in order to design effectively for trust, these nuances and complexities — the very things that make trust so unique and important — must be accounted for. Given the design gaps present, this research develops 'trust affordances' as a concept to deliver practicable insights into designing for trust.

1.2 Research Aims and Objectives

To address the overall aim of this research and the series of research questions developed in section 1.1, specific aims and objectives are developed to frame the research. These are presented in figure 1.2. As shown below, three specific aims correspond with the three research questions seeking to i) understand and ii) conceptualise trust within environmental data science, and then utilising these findings to iii) develop trust affordances to support the design of Virtual Labs. Two research objectives align to each of these specific aims, contributing to the completion of the research questions.

Research Question	**Specific Aim**	**Objective**	**Corresponding Chapter(s)**
1. How do stakeholders determine trustworthiness in order to place or refuse trust within environmental data science?	1. To understand how stakeholders determine trustworthiness in order to place or refuse trust within environmental data science.	1. Synthesise related trust literature to develop a suitable approach to studying trust.	2
		2. Identify and conduct an appropriate and suitable approach to research design.	4, 5, 6
2. What are the salient types of trust within environmental data science and how do they differ?	2. Conceptualise the nature of trust in environmental data science with respect to different dimensions of trust (people, data, and systems).	3. Analyse prevalent trust models in order to understand different trustors and trustees within environmental data science.	2
		4. Interview members of an environmental science Research Council and practicing environmental data scientists within a Centre of Excellence.	5, 6
3. How can considerations of trust be incorporated into system design?	3. Develop 'trust affordances' as a concept for design in fostering well-placed trust within Virtual Labs.	5. Theoretically develop and define trust affordances as a route to system design for fostering trust for multiple users.	3
		6. Re-analyse the empirical data with respect to trust affordances in order to illustrate their applicability.	5, 6

Figure 1.2: Research questions with corresponding research aims, objectives and chapters.

1.3 Research Contributions

The contributions of this research are three-fold. Firstly, this research provides rich empirical contributions to the understanding of trust within environmental data science, demonstrated by NERC's development strategy above. These contributions also extend to design, particularly the study of trust in Computer Supported Cooperative Work (CSCW) which focuses on supporting, facilitating, and coordinating interdependent cooperative and collaborative work (Bradner, Kellogg, and Erickson, 1999; Holtham, 1995; Rodden and Blair, 1991). The field of CSCW recognises trust as a fundamental ingredient within computing that requires further illumination (Cofta, 2007; Hill and Donaldson, 2015; Kelton, Fleischmann, and Wallace, 2008; Knowles, Rouncefield, et al., 2015; Marsh and Dibben, 2003). Whilst primarily focusing on Virtual Labs for future scientific collaboration and practice, greater knowledge is needed to inform their design particularly with regards to trust. In order to do this, this research also considers alternative technological opportunities for fostering trust, namely distributed ledger technology and data repositories, in order to determine potential trust affordances arising from Virtual Labs (Mackay et al., 1998; Olson and Olson, 2000; Sellen and Harper, 2001). This comparison enables a greater analytical insight, considering the specific elements one technology has versus another and how might these contribute to the formation of trust. For instance, Hine (2006, p.272) notes that databases are not new networks for disseminating science but that they "can provide the impetus for diverse new regimes to form" in comparison to traditional communications such as academic journals.

Secondly, rather than producing "a laundry list of facts and features" (Dourish, 2006, p.548) specific to one instance of a system in one context, this research introduces trust affordances enabling transferability and usablility to other systems and contexts (Dourish, 2001; Kaplan and Seebeck, 2001; Voss et al., 2009; Zimmerman and Nardi, 2006). Trust affordances are not limited to the research carried out here. In the spirit of CSCW, trust affordances also serve as an illustration

of system design as a whole, rather than its constituent parts (Dourish, 2001; Kaplan and Seebeck, 2001; Voss et al., 2009).

Finally, the development of trust affordances also point to another contribution: re-framing the process of system design. This research refutes a prevailing trend within the wider trust literature, namely that we must simplify trust in order to make it 'workable'. This research demonstrates that we do not have to ignore complexity or nuance, but rather, embrace it. Imbued with a sense of holism, this research serves as an illustration of the ways in which designers can creatively design for trust. This research presents trust affordances as an attempt to 'do things differently' whilst meaningfully representing trust in a socio-technical system comprised of broad, heterogeneous stakeholders.

1.4 Research Outline

In conjunction with the aims and objectives to address the research questions, figure 1.3 provides an outline of this thesis. The conceptual contributions of this research are presented in chapters 2 and 3 (indicated as 'theoretical' in figure 1.3). This includes the introduction of relevant trust literature, analysed through the application of two relevant trust models to the research aims, followed by analysis and synthesis within chapter 2. Chapter 3 builds upon the work produced by defining and developing the theory of trust affordances. In the methodological section (chapter 4), a qualitative case study research design is adopted and detailed to collect and analyse empirical data suited to the research at hand. In the empirical section (chapters 5 and 6) the data is analysed and re-analysed according to the research approach, producing an account of trust for different trustors and trustees across three main stakeholder groups within environmental data science. Each empirical chapter represents a separate case focusing on different stakeholder groups: data curators followed by data producers and consumers, and consideration of two technologies: blockchain and Virtual Labs, respectively. Following this, chapter 7

presents a high-level discussion on the research, building upon the reflections made in each chapter and conducting between-case analysis (with within-case analysis conducted within the respective empirical chapters). The questions, specific aims, and objectives developed within this introductory chapter will be assessed for level of completion and fitness of purpose in addressing the overall aim of this research in chapter 7. Finally, chapter 8 presents the conclusions and opportunities for future research.

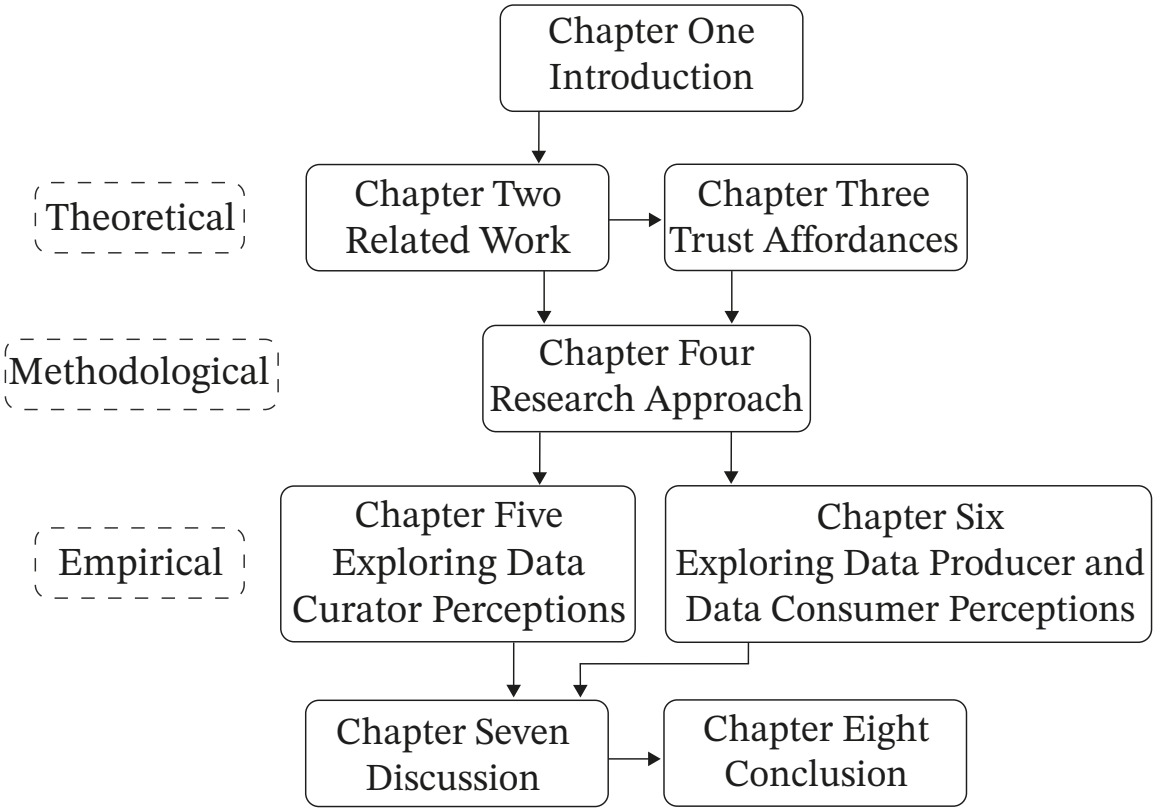


Figure 1.3: Structure of the thesis corresponding to the chapter content (theoretical, methodological, and empirical) in dotted lines on the left.

Chapter 2

Related Work

The practice of environmental data science contains inherent uncertainties, vulnerabilities, and risk, signaling that trust may be needed variously: at various times, for various people, for various reasons (chapter 1). To address this complexity, this research draws upon computing scholarship concerned with the interaction of users and technology, including HCI (Human-Computer Interaction) and CSCW (Computer Supported Cooperative Work). This subset places a focus on understanding users and their contexts as central to socio-technical design (Bannon, 2011; Norman and Draper, 1986; Schmidt and Bannon, 1992). Trust often features implicitly (and sometimes explicitly), given the ever-present interaction between humans and technology (Marsh, Atele-Williams, et al., 2020) and has developed an understanding of the factors that influence the trust placed: from credibility, usability, and risk (Corritore, Kracher, and Wiedenbeck, 2003) to levels of trust including technological, social, and institutional (Gach and Brubaker, 2020; Sas and Khairuddin, 2015). Yet there lacks a central focus upon trust as a defined, standalone subset of computing research. The theoretical portion of this research (figure 1.3) in both the present and succeeding chapters, draws upon two aspects of the computing literature, namely, user interaction and affordances and combines this with relevant trust-in-computing literature and beyond. Figure 2.1 illustrates key aspects from these two areas and how they intersect to inform the research.

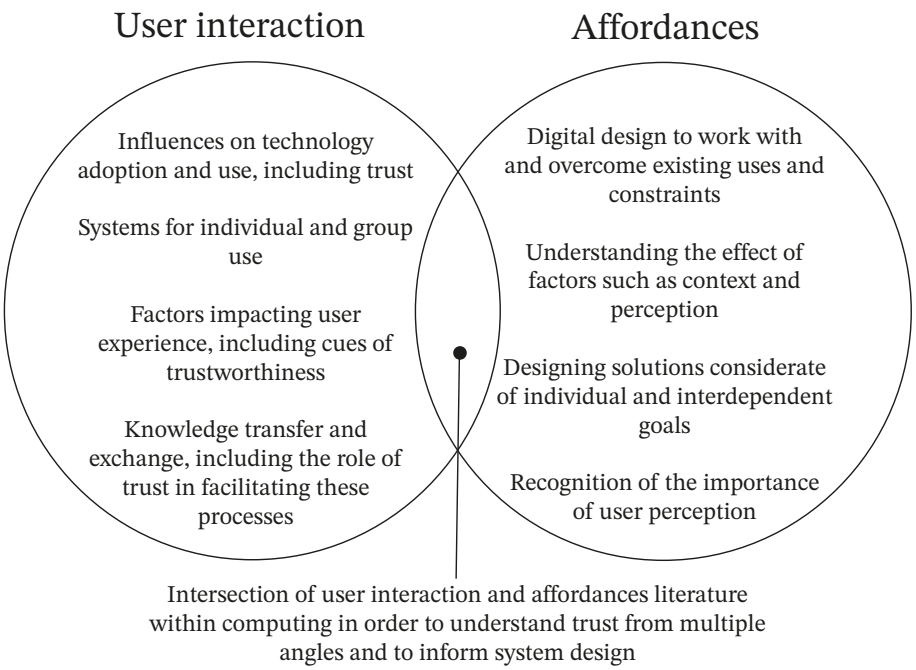


Figure 2.1: How the subsets of computing (user interaction and affordances) combine to inform the theoretical portion of the research.

As a construct with differing theoretical interpretations and hypothesised components, conceptual models are utilised to illustrate purported functions in differing contexts and to structure further research, such as system design (Patrick, Briggs, and Marsh, 2005). Consequently, models of trust as forms of hypotheses are a common feature of the cross-discipline literature on trust. However despite the plethora of conceptual trust models available for one to choose from, it remains challenging to select a model for a specific context, particularly for the research at hand which seeks to explore trust within transdisciplinary science with the aim of informing socio-technical system design. Thus in order to achieve a theoretical grasp of trust and how it may function (research question one, section 1.2), exploring trust models are of benefit but require prior effort in the form of appropriate selection before this can be undertaken. In order to study conceptual models of trust that resonate with this research, three aspects of the study of trust are first considered (section 2.1). This leads to the identification of two well-cited trust models that

are the most suitable for further theoretical exploration, conducted in sections 2.2 and 2.3. These sections jointly synthesise disparate streams of literature (trust, environmental science, and technological design) to provide a foundation for the empirical portion of this research and highlight dimensions of trust whilst developing a picture of how trust might function in this context. This chapter culminates in an assessment on the uses of trust models for system design via application (section 2.4), extrapolation (section 2.5), and reflection (section 2.6).

2.1 Approach to Trust

This research adopts a **social-psychological approach** to trust. An outline of this approach jointly: positions the research within the wider body of trust literature and conveys the implications of the approach, including how trust is *defined* (section 2.1.1), *studied* (section 2.1.2), and hence *modelled* (section 2.1.3). In sum, explicating the approach provides a basis from which to select models to analyse. This is particularly useful given the array of trust models proposed within the literature.

Social-psychological approaches to trust are associated with sociological, philosophical, and psychological disciplines, and often contrasted with an economic approach to trust (Cook and Santana, 2020; Lankton, Mcknight, and Tripp, 2015; McKnight, Choudhury, and Kacmar, 2002; Uslaner, 2018). Social-psychological approaches view trust as a mix of both emotion and cognition, and therefore as partly non-rational (Clark, 2014; Lahno, 2001b; Möllering, 2001; Möllering, 2006). Within this approach, trust aids navigation through uncertainty and is unneeded when we are certain, if we had the complete information to make a decision, this *would not be trust*, but rather confidence or reliability (Castelfranchi and Falcone, 2016; Luhmann, 2017; Mayer, Davis, and Schoorman, 1995; Simmel, 1950). In comparison, economics-based trust approaches posit that trust is a cost-benefit calculation, where the trustor is assumed to assess the costs and benefits before

making a rational, logical decision on whether to place trust (Coleman, 1990; Gambetta, 1988; Lewicki and Bunker, 1996; Nickel, Franssen, and Kroes, 2010).

Trust has been studied within CSCW from a social-psychological perspective (Benbasat and Wang, 2005; Corritore, Kracher, and Wiedenbeck, 2003; Knowles, Rouncefield, et al., 2015; Lankton, Mcknight, and Tripp, 2015; McKnight, Carter, et al., 2011; Passi and Jackson, 2018). Additionally, some studies have considered virtual research environment development (Schiff, Van House, and Butler, 1997; Van House, Butler, Ogle, et al., 1996) and collaborative systems (Jirotka et al., 2005) from a social-psychological perspective. This research builds upon this work, with the following subsections providing greater detail upon the approach to trust undertaken.

2.1.1 Defining Trust

Trust has been defined variously within and across disciplines (Blomqvist, 1997; Clément, 2020; Lewicki and Bunker, 1996), resulting in often conflicting definitions (Corritore, Kracher, and Wiedenbeck, 2003; Harper, 2014; McEvily, Perrone, and Zaheer, 2003; Patrick, Briggs, and Marsh, 2005). The choice of an approach, such as the social-psychological approach, does not lend itself to a specific, singular definition. A working definition that jointly accords with the research (Jantz and Giarlo, 2007; Patrick, Briggs, and Marsh, 2005; Sexton et al., 2005) and the multiple dimensions of trust within that research (McKnight, Choudhury, and Kacmar, 2002; Simon, 2020) must be selected.[1] The definition chosen for this research is adopted from Marsh and Dibben (2003, p.470), created directly for information and communication technology (ICT) using a social-psychological approach: "**trust concerns a positive expectation regarding the behavior of somebody or something in a situation that entails risk to the trusting party**".

Given the focus of this research, a definition which includes both human

[1]This definition will be reflected upon for alignment with participants and overall fit with the research in chapter 7.

and technological artefacts as members of a trust relationship is of paramount importance, demonstrated in the above definition by the inclusion of non-human trustees: "somebody or something". Moreover, this definition is commensurate with other independent definitions of trust, including the presence of vulnerability and uncertainty (Mayer, Davis, and Schoorman, 1995; Newell and Swan, 2000; Rousseau et al., 1998), signaling the dependability and credibility of the working definition adopted.

2.1.2 Studying Trust

Beyond a definition of trust, the social-psychological approach also highlights the inclusion of additional factors which guide the empirical study of trust, and importantly for this chapter, the choice of trust models. These factors are:

- **Trustors' Perceptions:** Within social-psychological approaches, trust is dependent on *trustors' perceptions*, in comparison to economics-based approaches wherein a focus is placed on the information available to the trustor (Benbasat and Wang, 2005; Blomqvist, 1997; Lee and See, 2004; Lewicki and Bunker, 1996; McKnight, Carter, et al., 2011). Thus any model chosen needs to focus on trustors' perceptions of trustees rather than the attributes of the information given to them.

- **Beyond Cognition:** Trust is thought of as a cognitive, considered process where a trustor deliberates all available information within economic-based approaches, but within social-psychological approaches is hypothesised to be based upon: affective or cognitive (Lewis and Weigert, 1985; Scheman, 2020; Simpson, 2012), heuristic or considered (Knowles, Smith-Renner, et al., 2018; O'Neill, 2020), mental or social processes (Lucas, Leith, and Davison, 2015; Möllering, 2013; Sabater and Sierra, 2005). Trust may be any of these processes or *a combination* thereof, dependent upon the trustor (Dwyer and Marsh, 2016; Riegelsberger, Sasse, and McCarthy, 2003b; Robbins, 2016);

the type of relationship between the trustor and trustee (Mayer, Davis, and Schoorman, 1995; Möllering, 2013); and the context of trust, e.g., where cognitive assessment outweighs affective inputs in a high-risk complex scenario (Castelfranchi and Falcone, 2016; Lewis and Weigert, 1985; Patrick, Briggs, and Marsh, 2005; Simpson, 2014). Thus any model needs to include factors beyond cognitive aspects of trust.

- **Inclusion of Context:** Within the social-psychological approach, the placement of trust is not solely about the perceived trustworthiness of the trustee, but the *surrounding context* (Kelton, Fleischmann, and Wallace, 2008; Kroeger, 2019; Misztal, 2020; Möllering, 2008). Trust develops within — and is influenced by — individual, cultural, and organisational contexts (Lee and See, 2004), e.g., an organisational context can influence how we perceive technology or data (Lee and See, 2004; McKnight, Carter, et al., 2011; Møller et al., 2020). Models must be selected that can account for context at varying levels, in order to capture the richness and dynamism of trust.

- **Reflection of the Multi-Faceted Nature of Trust:** It is recognised within the social-psychological approach that trust is always embedded in a social context and is thus different in practice for every person (Dwyer and Marsh, 2016; Möllering, 2005; Möllering, 2006). Approaches to determining trustworthiness and arriving at well-informed trust are idiosyncratic and 'multi-layered' (D'Cruz, 2018; Dwyer and Marsh, 2016). Any two individual trustors may respond or behave differently to the same set of inputs when placing trust (Castelfranchi and Falcone, 2016; Lee and See, 2004; McKnight, Choudhury, and Kacmar, 2002; Messick and Kramer, 2001; O'Neill, 2002; Riegelsberger, Sasse, and McCarthy, 2003b). Any trust models chosen must reflect this multi-dimensional and multi-faceted nature of trust across trustors, trustees, and contexts (Castaldo, Premazzi, and Zerbini, 2010; Corritore, Kracher, and Wiedenbeck, 2003; Ess, 2020; Lewicki, McAllister, and Bies,

1998; Nissenbaum, 2001; Rousseau et al., 1998) and the various levels (individuals, social groups, and institutions) in which trustors can place trust in trustees (Bierhoff and Vornefeld, 2016; McKnight, Choudhury, and Kacmar, 2002; Misztal, 2020; Zaheer, McEvily, and Perrone, 1998).

These factors thus point towards the suitability of existing trust models for application to the research goal, developed within the subsequent subsection, and in the empirical investigation of trust (chapter 4).

2.1.3 Modelling Trust

As detailed above, the approach to trust taken influences — but does not dictate — the research undertaken, particularly the definition, research focus, and modelling of trust. This subsection highlights the trust models chosen for the research, accounting for the factors highlighted in section 2.1.2 and additional criteria for trust models. These additional criteria include: flexibility to account for the different understandings and perspectives of trustors and the ways in which trust is placed; ability to model the various levels (individuals, social groups, and institutions) in which trustors can place trust in trustees (Bierhoff and Vornefeld, 2016; McKnight, Choudhury, and Kacmar, 2002; Misztal, 2020; Zaheer, McEvily, and Perrone, 1998); and inclusion of the different trust relationships at play with users and Virtual Labs.

Trust is relational, emanating from the relationship between trustor and trustee, and the environment (Cook and Santana, 2020; McKnight, Choudhury, and Kacmar, 2002; Sundqvist, 2011). This results in different ways in which trust is placed, illustrated in figure 2.2. Interpersonal trust traditionally occurs face-to-face but is increasingly mediated through technology, thus presenting a different set of challenges (Markus, 2001; Riegelsberger, 2002). Technological systems can be modelled as intermediaries of trust between users, e.g., trusted to provide secure computer-mediated communication (Castelfranchi and Falcone, 2016; Gach and Brubaker, 2020; Jirotka et al., 2005; Lankton, Mcknight, and Tripp, 2015). In addition to mediated trust through technology, there are also models of trust in

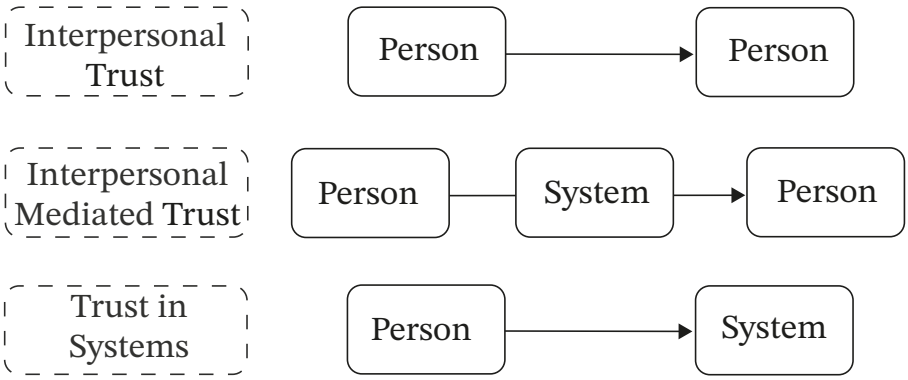


Figure 2.2: The different trust relationships present within the research area, where trust flows: from person to person linearly; is mediated between people by technological systems; and is placed in systems, respectively.

technology in which technological systems and artefacts are trustees with trust placed into them (Corritore, Kracher, and Wiedenbeck, 2003; Gulati, Sousa, and Lamas, 2017; Sas and Khairuddin, 2017; Söllner et al., 2012).

There are few models to choose 'off-the-shelf' for this specific application area (Thornton, Knowles, and Blair, 2022). Deciding which role a technology will take can impact on design choices and trust models (Lankton, Mcknight, and Tripp, 2015; Söllner et al., 2012). Virtual Labs can potentially be the objects of trust themselves (trustee); contain data and models which can also be objects of trust; and can, when designed for, enable communication between users (mediator). Thus two different trust models are chosen to represent these different trust relationships. The model developed by Riegelsberger et al. (Flechais, Riegelsberger, and Sasse, 2005; Riegelsberger, Sasse, and McCarthy, 2003b; Riegelsberger, Sasse, and McCarthy, 2005; Riegelsberger and Sasse, 2010) is chosen to model the system-as-a-mediator trust relationship ('mediator role trust model') and the model developed by McKnight, Carter, et al. (2011) is chosen to represent the system-as-a-trustee relationship ('trustee role trust model'). These models are shown in figure 2.3 as specified by their authors, and are addressed in sections 2.2 and 2.3 respectively.

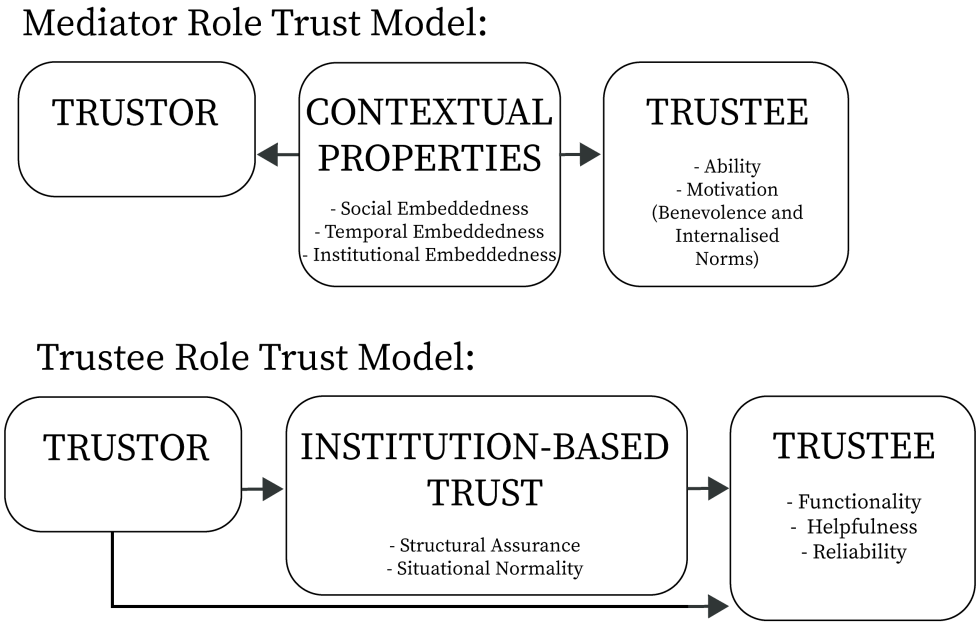


Figure 2.3: Mediator role model based upon the model proposed by Riegelsberger et al. (Riegelsberger, Sasse, and McCarthy, 2003b, Fig. 1, p. 771) and (Riegelsberger, Sasse, and McCarthy, 2005, Fig. 5, p. 399) (top) and trustee role model based on McKnight et al. (McKnight, Carter, et al., 2011, Fig. 2, p. 12) (bottom).

2.2 Modelling the Mediator Role

The *mediator role* focuses on trust between users mediated by technology (Söllner et al., 2012) resembling interpersonal trust in an online environment (figure 2.2). The model by Riegelsberger et al. (Flechais, Riegelsberger, and Sasse, 2005; Riegelsberger, Sasse, and McCarthy, 2003b; Riegelsberger, Sasse, and McCarthy, 2005; Riegelsberger and Sasse, 2010) (depicted on the top row of figure 2.3) utilises the ABI model of interpersonal trust within organisations, which postulates that there are three primary perceived factors of trustworthiness: ability, benevolence, and integrity (Mayer, Davis, and Schoorman, 1995). This model was chosen for the research context because of its applicability to trust in people and trust in technological systems. Regarding the latter, Riegelsberger, Sasse, and McCarthy (2005) state that because of the wider socio-technical context in which

a technological system sits, the model can be applied to trust in technology, as demonstrated by the application of the model in designing for trust within Human-Computer Interaction (HCI) by Sas and Khairuddin (2017).

The trust model developed by Riegelsberger et al. (Flechais, Riegelsberger, and Sasse, 2005; Riegelsberger, Sasse, and McCarthy, 2003b; Riegelsberger, Sasse, and McCarthy, 2005; Riegelsberger and Sasse, 2010) combines two *intrinsic properties* of a trustee — 'ability' and 'motivation' — alongside *contextual properties* (motivation based on temporal, social, and institutional embeddedness) proposing that in combination these properties form the basis of trustworthy behaviour (Flechais, Riegelsberger, and Sasse, 2005; Riegelsberger, Sasse, and McCarthy, 2005). In what follows, the model is further examined, considering both the trustee and contextual properties made by the Riegelsberger et al. in order to provide an in-depth treatment of individual model components.

2.2.1 Trustee Properties

Trustee properties are intrinsic and relatively stable attributes of the person or system in which trust is placed.[2] *Ability* has domain-specific (e.g., technical knowledge) and general components (e.g., intelligence) (Riegelsberger, Sasse, and McCarthy, 2005). This mirrors Mayer et al.'s ability characteristic which pertains to the "skills, competencies, and characteristics that enable a party to have influence within some specific domain" (Mayer, Davis, and Schoorman, 1995, p. 717). Ability is domain-specific, a trustee may have competence or the skills necessary to complete specific tasks, but this competency or skillset does not automatically indicate completion of a separate task for which they "may have little aptitude, training, or experience" (Mayer, Davis, and Schoorman, 1995, p. 717). In addition to trust in people, ability can be an antecedent for trust in systems and trust in data.

[2]Intrinsic properties can also relate to the trustor, e.g., propensity to trust and take risks (Flechais, Riegelsberger, and Sasse, 2005; Riegelsberger, Sasse, and McCarthy, 2005; Riegelsberger and Sasse, 2010).

Riegelsberger, Sasse, and McCarthy (2005) state that ability can be applied to trust in technology, relating to expertise, predictability, credibility, accuracy, authenticity, and availability. Ability may be perceived in systems through properties such as confidentiality, integrity (accuracy and reliability), authentication, non-repudiation, access control, and availability (Riegelsberger and Sasse, 2010). Regarding trust in data, authenticity, accuracy and credibility of data are important characteristics in determining whether to place trust (Faniel, Kriesberg, and Yakel, 2016; Fear and Donaldson, 2012; Knowles, Harding, et al., 2014; Sexton et al., 2005), and to those who created it as trust in data is derived in part from trust in data producers (Metzger and Flanagin, 2013; Van House, 2002; Yoon, 2014).

Motivation is divided into benevolence (related to affective trust) and internalised norms (related to cognitive trust) (Riegelsberger, Sasse, and McCarthy, 2003b). *Benevolence* reflects "the extent to which a trustee is believed to want to do good to the trustor, aside from an egocentric profit motive" (Mayer, Davis, and Schoorman, 1995, p. 718). Benevolence is only typically related to forms of continuous interpersonal trust between a trustor and a trustee, however, Riegelsberger et al. note that there can be another type of benevolence, from organisations towards consumers (Riegelsberger, Sasse, and McCarthy, 2005). If we define a virtual research environment as connected to and built by a specific institution or organisation, then benevolence could be appropriate concerning trust in systems, e.g., by signaling the ways in which they go above and beyond to help users with their concerns; expressing their commitments to trustworthy scientific research; or dependent on system builders. Additionally, benevolence can also be connected to data producers, who act benevolently by providing their data to others via repositories, despite the lack of a continuous trust relationship between a data producer and consumer.

Internalised norms (also referred to as 'dependability') refers to the perception of a trustee acting in line with the principles, internalised codes of conduct, norms and values that they purport to uphold (Riegelsberger, Sasse, and McCarthy, 2003b; Riegelsberger, Sasse, and McCarthy, 2005), bearing similarity to 'integrity' (Mayer,

Davis, and Schoorman, 1995). Internalised norms are important for trust in people, e.g., honesty and value congruence. Internalised norms could be applied to the virtual research environment wherein responsiveness, openness, good will, principles, values, and standards are all connected to internalised norms (Riegelsberger, Sasse, and McCarthy, 2005). Yoon (2014) found that 'organisational attributes' (including integrity and honesty, commitment to users and society, and values) was an important source of trust for repository users in the repository as a system. The trust in a repository as a system connects to trust in data, as repositories are trusted sources for data (Yoon, 2014). Whilst data itself may not have internalised norms, these can be inferred from either the people or systems it comes from, for instance honesty is used to determine credibility (Corritore, Kracher, and Wiedenbeck, 2003).

2.2.2 Contextual Properties

Building upon Mayer et al., who noted that "the specific consequences of trust will be determined by contextual factors such as the stakes involved" (Mayer, Davis, and Schoorman, 1995, p. 726), Riegelsberger et al. include three *contextual properties* within their model (figure 2.3). Contextual properties provide both positive and negative incentives for trustors and trustees as they can motivate trustees to act in a trustworthy manner and can support the placement of trust for trustors (Riegelsberger, Sasse, and McCarthy, 2005; Riegelsberger and Sasse, 2010).

Temporal embeddedness occurs if there are stable identities, and the trustee has reason to believe that they will interact again in the future and be recognisable (Riegelsberger, Sasse, and McCarthy, 2005; Riegelsberger and Sasse, 2010). Membership of a community or organisation can support trustworthy behaviour and is 'institutionally assured' because of the likelihood of further encounters (Cook and Santana, 2020; Riegelsberger, Sasse, and McCarthy, 2005). If users are identifiable, it is likely that they want to maintain their reputation and so are encouraged to submit trustworthy data and to act accordingly within the system itself thus aiding trust in data. Temporal embeddedness can aid trust in the system and data,

particularly in terms of longevity wherein a repository may be trusted to preserve data reliably and accurately in the long-term (Yoon, 2014).

Social embeddedness refers to a trustee's performance and reputation regarding their honesty, reliability, or dependability within a community and can provide an incentive to fulfil, even without prospect of specific future interaction (Riegelsberger, Sasse, and McCarthy, 2005). There are real-world risks, e.g., embarrassment, to being or appearing to be untrustworthy in front of others, but when a trustee demonstrates trustworthiness there are rewards, displaying their trustworthiness to others that bear witness to this (D'Cruz, 2019; Pettit, 1995; Riegelsberger, Sasse, and McCarthy, 2003b; Yakel et al., 2013). Even without the possibility of future direct interaction between a data producer and a data consumer, reputation is precious, being built over time through sustained effort, and can be 'lost' (Meyerson, Weick, and Kramer, 1996). Social embeddedness can also relate to the credibility of online information itself and the data producer responsible for its creation, e.g., honesty, expertise, predictability, and reputation (Corritore, Kracher, and Wiedenbeck, 2003; Passi and Jackson, 2018; Zimmerman, 2008a), where feelings of (dis)trust in data are related to feelings of (dis)trust in the organisation or institution responsible for the data (Steedman, Kennedy, and Jones, 2020).

Finally, *institutional embeddedness* considers the influence of institutions or organisations, including wider networks and third parties (Riegelsberger, Sasse, and McCarthy, 2005; Riegelsberger and Sasse, 2010). These institutions or organisations can directly influence the behaviour of their members, through codes of practice, the monitoring of standards, and the threat of being expelled (Bovens, 2007). Institutional embeddedness can signal information about the intrinsic properties of a trustee, e.g., their ability or honesty, or their professional qualifications, e.g., through rigorous testing that results in accreditation (Riegelsberger, Sasse, and McCarthy, 2005). As a form of institutional trust, this signals trust in people and their data, because it acts as an indicator of trustworthiness but does not rely on interpersonal familiarity with a trustee. For these reasons, institutional embeddedness can also

assure trust in data through standards and rules of conduct.

2.3 Modelling the Trustee Role

The model by McKnight, Carter, et al. (2011) was chosen to model the user-system relationship where the technology is the *trustee* (Corritore, Kracher, and Wiedenbeck, 2003; Söllner et al., 2012). This model, illustrated in the bottom row of figure 2.3, consists of: propensity to trust general technology (consisting of 'trusting stance' and 'faith in humanity'), institution-based trust in technology ('structural assurance' and 'situational normality'), and trust in a specific technology ('functionality', 'helpfulness' and 'reliability') with a causal ordering where propensity to trust influences institution-based trust and indirectly influences trust in a specific technology (McKnight, Carter, et al., 2011). In the same manner as the mediator role model, the trustee and contextual components of the model will be examined in order to provide sufficient detail of the model.

2.3.1 Trustee Properties

Trusting beliefs in a specific technology are trustee-specific and based on a relationship with a particular technology. These beliefs are related to the ABI attributes developed by Mayer, Davis, and Schoorman (1995) but "are less likely to violate humans' understanding of a technology's capabilities" (Lankton, Mcknight, and Tripp, 2015, p. 883). *Functionality* refers to the belief that the technology has the functionality and capability to complete a specific task. Thus, many of the applications of 'ability' from the mediator role above are valid in the trustee role model.

Helpfulness refers to the belief that the specific technology provides adequate, effective, and responsive help for users when needed, bearing similarity to 'benevolence' (Mayer, Davis, and Schoorman, 1995) but excluding moral agency and volition (McKnight, Carter, et al., 2011). Regarding both trust in data and trust in systems,

helpfulness may be realised in various aspects of system design: help functions, connection to staff, or communication with other Virtual Labs *if* these resources are designed into a system.

Finally, *reliability* refers to continuous, reliable operation and predictable response (McKnight, Carter, et al., 2011). Due to the focus exclusively upon trust in systems within this model, the interpretation of 'integrity' does not relate to commitments and principles but rather system functionality and the absence of failures and downtime. A system that does not work and is therefore unusable would not enable to users to develop initial trust. This trustee attribute could also be interpreted in the form of longevity and continuous trust, which can take a long time to foster between users and systems (Knowles, Rouncefield, et al., 2015). This is important in the introduction of a new system such as a Virtual Lab where there is uncertainty in the long-term provision of services and data (Lin et al., 2020). Reliability as defined within the trustee role model bears a resemblance to 'temporal embeddedness' within the mediator model, particularly the "absence of technical failures, absence of mistakes, breadth of product palette, aesthetic design, and information about physical assets" (Riegelsberger, Sasse, and McCarthy, 2005, p.410).

2.3.2 Contextual Properties

Institution-based trust consists of *technological structural assurance* and *technological situational normality*. These constructs are adapted by McKnight, Carter, et al. (2011) for trust in technology. *Structural assurance* refers to the belief that structures such as guarantees and measures (legal and technical) are in place (McKnight, Choudhury, and Kacmar, 2002; McKnight, Carter, et al., 2011; Shapiro, 1987; Zucker, 1986). Where technology is a trustee, this implies that the structural conditions exist to make success with the technology likely regardless of the characteristics of the specific technology (McKnight, Carter, et al., 2011). This may be perceived when guarantees, contracts, and support are in place with the

technology (McKnight, Carter, et al., 2011). These are more likely in commercial or organisational settings rather than academia. However, support may be possible if the virtual research environment is affiliated with an institution or organisation, thus providing a guarantee through institutional association, for instance.

Technological situational normality relates to the belief that the system is functioning in a predictable, normal and well-ordered way so that a trustor can extend trust to a new trustee with a successful outcome, i.e., well-placed trust (Baier, 1986; Lewis and Weigert, 1985; Luhmann, 2017; McKnight, Choudhury, and Kacmar, 2002; McKnight, Carter, et al., 2011). This is relevant if users have experience with virtual research environments (or similar technologies) hence impacting on their belief in success. However, as noted previously (chapter 1), Virtual Labs are a relatively new class of technology with few implementations that users may draw upon in terms of experiential knowledge. This may be supplemented with a history of using data repostiories and insitutional affiliation of Virtual Labs, as is the case with Data Labs developed by UKCEH (Hollaway et al., 2020).[3]

2.4 Application to the Research

Several models were considered before opting for the models by Riegelsberger et al. and McKnight et al. For instance, the model of 'trust in automation' by Lee, Moray, and See (Lee and Moray, 1992; Lee and See, 2004; Muir and Moray, 1996) is also based upon the ABI model (Mayer, Davis, and Schoorman, 1995). However as automation is defined as "technology that actively selects data, transforms information, makes decisions, or controls processes" (Lee and See, 2004, p. 50) their model is not directly applicable to this research goal. The two chosen models represent a social-psychological approach to trust, including trustors' perceptions, affective trust, and context (section 2.1.2). As Virtual Labs could be perceived either as a trustee for some stakeholders or as a mediator by others, these two models

[3] `https://datalab.datalabs.ceh.ac.uk` accessed July 2023.

were chosen as the most applicable for each role. Purely interpersonal models of trust would not have been able to capture the landscape of this research: a socio-technical design requires socio-technical trust models. And, despite the variety and volume of models within the wider trust literature, there are few models tailored to the socio-technical study of trust. The models use similar characteristics, both adapting the three antecedents of trust by Mayer, Davis, and Schoorman (1995) to their own application areas: the mediator role model by Riegelsberger et al. takes an interpersonal trust route, whereas the trustee model by McKnight et al. takes a trust in technology route. Given that perception of trust in a system varies, Sommerville et al. (2006) posit that a system within an organisation is more likely to be perceived as trustworthy based upon 'trustee' characteristics such as dependability for organisational members, but for external stakeholders, this degree of trust varies according to previous experience, reputation of the organisation, and the availability of recourse and redress if failures occur. At this conceptual stage, it is unknown how a system is perceived within environmental data science across the spectrum of users, warranting the choice of two similar but different models. These models were the most suited to Virtual Labs and the research approach to trust, as they are cognizant of the multi-faceted nature of trust, the importance of context, and that trust is based upon a trustors' perceptions of both emotive and cognitive cues (section 2.1.2). Overall, the models enabled a theoretical exploration of the trust space within the socio-technical design of a Virtual Lab, providing a basis for both thinking about and analysing trust. Without models of trust, there would be no means to understand and importantly contrast lines of thought between authors. Moreover, they are a useful means to highlight potential complexities within trust that warrant further exploration. The plethora of models within the trust literature speaks to their ability to communicate what is a truly complex phenomena.

One of the key contributions from this process is the dissection of these models' components, presented in figure 2.4, wherein trust in people, data, and systems are plotted to the relevant model properties. Many of the trustee and contextual

	Institutional Characteristics	Trustee Characteristics
Mediator Role (Riegelsberger et al.)	*Temporal Embeddedness (TD, TP, TS)*	*Ability (TD, TP, TS)*
	Social Embeddedness (TD, TP)	*Motivation - Benevolence (TP)*
	Institutional Embeddedness (TD, TP, TS)	*Motivation - Internalised Norms (TP, TS)*
Trustee Role (McKnight et al.)	*Situational Normality (TD, TS)*	*Functionality (TS)*
	Structural Assurance (TS)	*Helpfulness*
		Reliability (TP, TD, TS)

Key: TD (trust in data), TP (trust in people), TS (trust in systems)

Figure 2.4: Application of trustee and mediator role models to the design of a transdisciplinary virtual research environment (Thornton, Knowles, and Blair, 2022, p.1392).

properties were relevant to either or all trusts under consideration. Some findings, such as the wide applicability of the mediator model in comparison to the trustee model are expected, given the purpose of the model (trust in people and socio-technical systems vs. trust in systems respectively). Other properties such as 'helpfulness' were not connected to any trust dimension. This is due in part to the definition presented by McKnight, Carter, et al. (2011), wherein the characteristic is described by the presence of a system help function. However, this could be applicable if defined differently, i.e., from other users in the Virtual Lab providing help and support. Gulati, Sousa, and Lamas (2019) argue that helpfulness in terms of adequacy, effectiveness, and responsivity of a computer system would fall under 'benevolence', suggesting overlap between the two models.

The mediator role subsumed many of the properties from the trustee role model, e.g., structural assurance and situational normality into institutional embeddedness. Again, the definitions by McKnight et al. of these contextual properties influenced their application reflected in figure 2.4. 'Institutional embeddedness' was connected to trust in people, data, and systems but 'situational normality' and 'structural assurance' were connected to trust in data and systems and trust in systems, respectively. The contextual properties (three forms of embeddedness) from

Riegelsberger et al. provided a real grounding for socio-technical system design, and going further beyond Mayer et al., and later Schoorman et al., who noted the importance of context, but did not explicitly model it.[4] Firstly understanding, and then modelling, context is imperative and should not be overlooked.

Given their social-psychological stance, the models also reflect the belief that trust is not purely cognitive and that elements of interpersonal and institutional trust are important. Despite this, whilst affective trust is featured, cognitive trust is the most prevalent. This is illustrated by the lack of applicability of 'benevolence' and 'helpfulness' in comparison to 'ability', 'internalised norms' and 'functionality' (figure 2.4).[5] In the same vein, both models recognise that trust functions across levels as interpersonal and impersonal trust (Fulmer and Dirks, 2018; Schilke and Cook, 2013; Shapiro, 1987). Institutional trust is a key part of the social-psychological approach (McKnight, Choudhury, and Kacmar, 2002; Misztal, 2020; Zaheer, McEvily, and Perrone, 1998), seen as a supportive pillar of trust formation in both people (Blomqvist, 1997; Kroeger, 2019; Lahno, 2001a; Sitkin and Roth, 1993; Zaheer, McEvily, and Perrone, 1998) and technology (Button, 2006; Castelfranchi and Falcone, 2020; Konrad, Fuchs, and Barthel, 1999). Trust is placed in the function of the system *circumventing the need for direct experience* or a *complete understanding* of the trustee (Giddens, 1990; Konrad, Fuchs, and Barthel, 1999; Lewis and Weigert, 1985; Luhmann, 2017). With regards to science as an institution (Latour and Woolgar, 1979; Oreskes, 2019), institutional trust can reduce the burden of assessing trustworthiness given its heritage in standards and practices (Blomqvist, 1997; Rolin, 2020; Zucker, 1986). Trust may be developed by virtue of a trustee being a member of a scientific community and the connotations that this brings (Fukuyama, 1995; Van House, Butler, and Schiff, 1998). Yet the role that this plays for both practicing scientists trusting in different scientific communities or for non-scientists trusting in the institution is not clear (Van House, 2002). Impersonal

[4]See: 'The Role of Context' in Mayer, Davis, and Schoorman (1995, pp.726-727) and 'Context-Specific Models of Trust' in Schoorman, Mayer, and Davis (2007, p.351).

[5]Though as described above this is due in part to the definition of characteristics by authors.

trust differs to other forms of trust (Marsh and Dibben, 2003; Rousseau et al., 1998; Simon, 2020) and thus *solutions to foster either of these will necessarily be quite different from one another*. A key consideration within this research, as posited within the first and second research questions, is to understand how trust is placed so that we can then determine how to design a system based around these findings.

Key Insights

- There are many possibilities for designing new technology and fostering trust for users, but there is no clear way to determine how these features or functions may (dis)satisfy users and how one should go about identifying and then organising these.
- Despite the abundance of trust models there is a paucity relating to socio-technical system design. Models are designed to be a general and approximate representation of a phenomena, thus lifting trust models between different scenarios will not yield the best results. However, analyses of models via deconstruction and examination, as demonstrated within this chapter, can be undertaken by system designers to apply pre-existing models to their context.
- This approach not only shows in what ways models fit a context, but also serves as a means to highlight how they underserve a context. Specifically, that trust in data – despite being a cornerstone of scientific research – has to-date not been modelled. Section 2.5 considers this further.
- A central argument within this research that has not been stated is that we cannot seek to design a technology to support trust without first deeply understanding trust. And yet, having sought to understand it, as demonstrated thus far, emergent designs are far from apparent. Section 2.6 reflects upon the role of models within system design.

2.5 Extrapolating Beyond the Trust Models

"The practice of data collection—even in science—is informed by disciplinary convention, situational contingencies, and individual flair" (Feinberg, 2017, p.2953). This is true of environmental data, containing many complexities, idiosyncrasies, and uncertainties (Bowker, 2000; Zimmerman, 2008b). For instance, temperature can be measured in many ways or transformed from other variables by satellites (Borgman et al., 2007; Edwards, 2010); sensors, often assumed to be inherently objective, can be biased depending on their placement and other contextual factors (Carlson and Anderson, 2007; Knowles, 2016; Wallis, Borgman, et al., 2007); and observational data, whether conducted by a citizen scientist or a trained scientist, is not completely objective (Birnholtz and Bietz, 2003; Dourish and Cruz, 2018; Van House, 2002). In this regard environmental data science is rare as a discipline, containing both small-scale measurements and national or multi-national process models. Whilst traditionally managing these challenges, the turn towards transdisciplinary data science introduces more complexities including higher spatial and temporal measurements, and a greater distance between the data consumer and the original context in which the data was collected or produced (Baker and Yarmey, 2009; Bates, Lin, and Goodale, 2016; Neff et al., 2017).

As noted in chapter 1, trust in data, people, and systems are interconnected (Jantz and Giarlo, 2007; Knowles, Rouncefield, et al., 2015; Knowles, 2016; Marsh, Atele-Williams, et al., 2020). Given the importance of data in science (Kratz and Strasser, 2015), developing an understanding of trust in data is paramount to design a system to facilitate well-placed trust (or distrust). As highlighted above, there is a scarcity of models concerning trust in data when compared to trust in people or trust in data, necessitating a deeper consideration and combination the work presented thus far within this chapter.[6] In what follows, the application of models conducted

[6]Notable related but not exact examples include trust in digital information (Kelton, Fleischmann, and Wallace, 2008); trust in informational or transactional websites (Corritore, Kracher, and Wiedenbeck, 2003); and trust in secondary data reuse (Yoon, 2017).

thus far is used as a starting point in order to develop a more comprehensive picture of trust in data.

As described in sections 2.2 and 2.3, trust in data relates directly to certain trustee characteristics: 'ability' from the mediator role model and 'reliability' from the trustee role model. This is reflected within the literature, wherein digital information may be evaluated in terms of reliability (Faniel and Jacobsen, 2010) or accuracy, objectivity, validity (soundness and verifiability) and stability (predictability and persistency) (Kelton, Fleischmann, and Wallace, 2008) Trust in data is also related to contextual properties: all three forms of embeddedness from Riegelsberger et al., and 'situational normality' from McKnight et al. (figure 2.4) and may take different forms, e.g., using provenance in order to determine data quality via data producer reputation and expertise (Fear and Donaldson, 2012). Beyond the initial connections made, there are also other *potential* connections to be made, illustrated in figure 2.5. For instance, trust in data may be determined through reputation and recommendation (hence mediated by trust in people through 'social embeddedness', also referred to as 'social trust' (Hill and Donaldson, 2015; Kelton, Fleischmann, and Wallace, 2008)). Faniel, Kriesberg, and Yakel (2016) suggest that reputation is used to evaluate trust and is based upon: lineage, institutional affiliation, competence, commitment, past performance, prior experience, and shared orientations and values. Others suggest that this information is not as important. For instance, Cragin and Shankar (2006) suggest that particulars about the data itself are more important than any information gleaned regarding the data producer, and Zimmerman (2008a, p.647) states that "knowing people and recognizing who does good work does not lead to automatic acceptance of data, but it does help to lessen concerns about data quality".

The cues of trustworthiness relevant to developing trust in data are many and varied (Borgman, 2007; Marsh, Atele-Williams, et al., 2020; Sexton et al., 2005). To illustrate these potential connections two aspects of trust in data not discussed within the literature are focused upon: 'helpfulness' (not directly connected to any

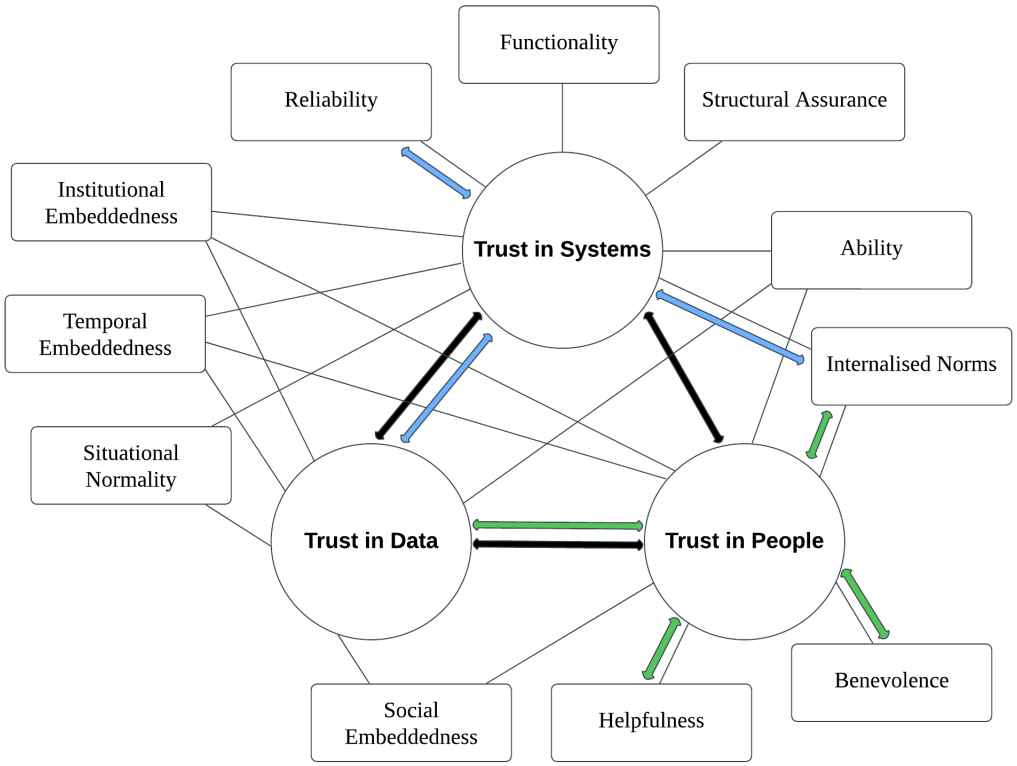


Figure 2.5: Connection between trust in data and model characteristics, and how trust in data is mediated by trust in people and trust in systems. Thick black arrows denote the interrelations between trust in: people, data, and systems. Thin black lines show the connections made in Thornton, Knowles, and Blair (2021). Green and blue lines illustrate how trust in data is mediated by trust in people and trust in systems respectively.

dimension of trust, figure 2.4) and internalised norms (mediated by people and by the system).

Helpfulness can be reformulated to consider whether the data helps the user to achieve their task, i.e., appropriateness (Faniel, Kriesberg, and Yakel, 2016; Yoon, 2017). Hind et al. (2020) state that the absence of useful information diminishes perceived trustworthiness. In this case, supplementary information ('meta-data') can be helpful (Borgman, 2007; Faniel and Jacobsen, 2010; Fear and Donaldson, 2012; Yoon, 2014). Consistent and useful documentation is difficult

to create, *especially* as the audience for data reuse does not necessarily emanate from that discipline and thus there is no general baseline of assumed knowledge or understanding of context (Bowker, 2000; Faniel, Frank, and Yakel, 2019; Schiff, Van House, and Butler, 1997). Where available, this information may be particularly important for certain data consumers (i.e., those lacking familiarity with disciplinary norms and conventions and with no knowledge to base reputation and affiliation) in order to develop trust. As with the role of reputation in forming trust, the literature shows diverging opinions: data users may use information such as meta-data to determine accuracy (Faniel, Frank, and Yakel, 2019), however, for some users, meta-data alone may not be sufficient enough to determine whether to place trust (Fear and Donaldson, 2012).

Regarding *internalised norms*, Kelton, Fleischmann, and Wallace (2008) suggest that information can be trusted based upon its perceived integrity and honesty. Trustworthiness requires transparency about imperfections (Faniel and Jacobsen, 2010; Knowles, 2016). This trust in information can be the data itself, for instance validity which includes "soundness of the methods used, the inclusion of verifiable data, and the appropriate citation of sources" (Kelton, Fleischmann, and Wallace, 2008, p.370), or alternatively, from the source (either data producer or data curator) (Corritore, Kracher, and Wiedenbeck, 2003; Fear and Donaldson, 2012; Marsh, Atele-Williams, et al., 2020; Yoon, 2014). Trust can stem from the policies and standards related to that specific system and/or the hardware and software in use (Jantz and Giarlo, 2007) or from reputation which acts as a 'trust marker' for data reuse (Faniel and Yakel, 2017). As discussed throughout this chapter, there are differing effects dependent on the trustor, with Yakel et al. (2013) finding that institutional reputation had a stronger effect on non-experts than experts.

In sum, despite a relevant model to represent trust in data within the literature, we can begin to see that there are kaleidoscopic patterns, with differing perspectives according to the user and their lens. This sets the grounding for empirical work to uncover the specific views of environmental data science stakeholders, their trust in

data, and whether this is perceived as trusting in the data itself and/or mediated by trust in people or trust in systems.

2.6 The Role of Trust Models in System Design

A key concern of this research is trust in environmental data (and scientific models), which involves both trust in technology and trust in people and/or institutions. This chapter reports upon the application of two conceptual trust models relevant to the design space, and an analysis and synthesis of the literature leading to a developed understanding of trust in data as not purely cognitive but including affective and social aspects, and that both interpersonal and impersonal trust play a role in forming trust. Importantly, this chapter highlights that there are many uncertainties and nuances left to uncover, signaling aspects which can be taken further within the research. Models can be useful when conceptualising a complex phenomenon and acting as a structure to further one's own understanding. Yet given the nature of trust, i.e., a nebulous and personal concept, it faces a similar treatment within academic literature. For this reason, there are many subtleties regarding trust, which neither existing models nor the literature have helped to overcome. Firstly, there are both data producers and data consumers to consider when designing a Virtual Lab. The literature focuses mostly upon data consumers, but there are also the perspectives of data producers to consider (Cragin and Shankar, 2006; Van House, Butler, and Schiff, 1998).

Secondly, these models don't reflect the flow of trust nor the level of familiarity existing between trustors and trustees. Uncertainty within science and with data is common (Beven, 2009; Pink, Lanzeni, and Horst, 2018) and scepticism is an accepted feature of science (Ramírez-i-Ollé, 2018). Investigating data prior to use does not point towards a lack of trust (Faniel, Kriesberg, and Yakel, 2016; O'Neill, 2018; Ramírez-i-Ollé, 2018), but rather shows that trust requires "intermittent information processing" (McEvily, Perrone, and Zaheer, 2003, p.99). This speaks

to a different form of trust between *initial trust* which occurs before interaction based on indirect, second-hand information, experience with similar situations, or cognitive cues (Barber, 1983; Li, Hess, and Valacich, 2008; McEvily, Perrone, and Zaheer, 2003; McKnight, Cummings, and Chervany, 1998; McKnight, Choudhury, and Kacmar, 2002; Meyerson, Weick, and Kramer, 1996) and *continuous trust* which can form based on a direct experience or familiarity with the trustee and repeated interaction over time (Hoehle, Huff, and Goode, 2012; Lewicki and Bunker, 1996; Rousseau et al., 1998). Within the general trust literature, it is suggested that if a trustor already trusts a trustee then they may ignore contrary evidence and forgo collection of any additional information (Frost-Arnold, 2020; Jones, 1996; Jones, 2019; McKnight, Cummings, and Chervany, 1998). This appears to be antithetical to the values and pursuits that science holds, as validation of data is commonplace.

Thirdly, this developed understanding does not provide a clear path towards the design of a Virtual Lab. As discussed in the introductory chapter, Virtual Labs are flexible in implementation and thus can be designed to elicit certain experiences through the physical properties and features of the system and of the Virtual Lab in itself. As documented within this chapter, utilising models to explore the concept of trust has been a worthwhile approach, enabling a deeper understanding of different components of trust and how they work together or conflict. However, there is no clear direction from models to design.

Ultimately, models seek to simplify and represent often complex phenomena and can arrange and organise a concept into a meaningful composition where none exist (Patrick, Briggs, and Marsh, 2005). Models are not expected to be perfect representations of reality (Marsh, Atele-Williams, et al., 2020; Michener, 2000). However, using them for the purpose of designing a system is challenging, i.e., trust models may only cover a subset of characteristics of interest (Donaldson and Conway, 2015); not sufficiently explain how trust is produced (Misztal, 2020); and may not reflect the multi-faceted nature of trust sufficiently (Castelfranchi and Falcone, 2016). Castelfranchi and Falcone (2010) and Schiff, Van House, and Butler (1997)

note that even in the models reviewed within this chapter, each characteristic can be too difficult to apply in practice with a lack of insight within each 'box' (i.e., the individual components of models). One needs to balance the need for a parsimonious model and creating a model that reflects the complexity of trust (Blomqvist, 1997; Marsh, Atele-Williams, et al., 2020; Schoorman, Mayer, and Davis, 2007).

As argued in Thornton, Knowles, and Blair (2022), models lack the malleability to characterise different types of trustors and trustees within different contexts sufficiently to map onto system design. Trust has "intricate contours" (D'Cruz, 2018, p.249) and as shown, models only go part of the way in attending to the socio-technical, i.e., the technical, social, and cultural factors in which the system sits (Anderson and Sharrock, 1992; Baxter and Sommerville, 2011; Davis, Challenger, et al., 2014). These surrounding contextual factors have effects on the development of trust (Blomqvist, 1997; Dwyer and Marsh, 2016; Kroeger, 2019), shaping perceptions of trustworthiness (Corritore, Kracher, and Wiedenbeck, 2003; Lee and See, 2004; Mayer, Davis, and Schoorman, 1995; Møller et al., 2020). Models are unable to create the distinction between the multiple requirements for Virtual Labs (Markus, 2001; Marsh, Atele-Williams, et al., 2020) and the potential competing interests of system users (Pirson and Malhotra, 2010; Sexton et al., 2005; Van House, Butler, and Schiff, 1998). In chapter 3 this theoretical work is taken forward using the experience of model application — and the alchemy needed to successfully design for trust — to develop the theory of trust affordances.

2.7 Chapter Takeaways

This chapter presents, analyses, and synthesises an array of related literature through the application of two prominent models of trust. This includes defining trust and considering the different dimensions of trust (interpersonal, technological systems, and data) for different trustors and trustees. This chapter also represents consideration of two prominent models of trust and the development of an

understanding of trust in data, before discussing the role of models in consolidating and communicating such a complex phenomenon but a limited ability in applying this to design. This leads to aspects that will be taken forward (figure 1.2), including the development of design principles in the form of trust affordances within the next chapter, and developing a suitable research approach given the approach to trust taken. Finally, this chapter analyses, synthesises, and reflects upon the work undertaken. Altogether this results in a thorough examination of trust and paves the way for constructing a research design that speaks to the aims of the research. In order to design technology for the well-placed trust of multiple stakeholders, one must first fully inspect the myriad ways in which trust is built. Without doing so, the risk is that — even with the best of intentions — the final system does not fully attend to any users in any meaningful way. Models, whilst often proposed as the outcome of research are in fact only one stepping stone in the design process. In the context of the wider research, this chapter reveals a much-needed broader consideration of trust beyond 'trust in data' in order to study and inform the design of a related socio-technical system.

Chapter 3

Trust Affordances

Building upon the last chapter, this chapter continues the theoretical portion of this research by developing the concept of *trust affordances*. Affordances as found in computing relate to user-centred design and the development of technological systems and features to 'afford', or make possible, certain actions (figure 2.1). This research contends that the same concept can be utilised in order to design for trust whilst accounting for the myriad ways in which trust is developed for different stakeholders.

An overview of the origin of affordances (section 3.1) and their introduction to computing (section 3.2) are first presented in order to provide clarity and background to the concept. Following this, trust affordances are defined and relevant conceptualisations within the affordances literature are introduced to aid in this development (section 3.3). These conceptualisations provide the means to move from physical to social affordances of technology and establish a deeper understanding of affordances. Finally in section 3.4, trust affordances are developed vis-à-vis three themes prevalent in within the social computing literature: *provenance*, *transparency*, and *social translucency*. Using extant literature to develop an understanding of these themes jointly illustrates the opportunity for reasoning about trust and provides a grounding which will be further explored with empirical data to understand trust within environmental data science in chapters 5 and 6.

3.1 Gibson's Affordances

Gibson (Gibson, 1977; Gibson, 1979) coined the term 'affordance' in his research on ecological psychology. To Gibson, an affordance is a relation (action possibility) between an environment (or object in said environment) and an organism (human or animal), independent of the organism's ability to perceive this possibility (Gibson, 1977; Gibson, 1979). Gibson's conception – focused on visual perception and the direct pick up of cues – was formed in opposition to earlier cognitive psychology which described perception as a result of information processing (Fayard and Weeks, 2014). An affordance is based on an individual's subjective perception of utility and the objective qualities of an environment or object (Gibson, 1979). The same environment may offer different affordances to different organisms dependent on their ability and activity, thus affordances can all be accessed in the same way, but only *some* animals can access the affordances (Gibson, 1979; Kaptelinin and Nardi, 2012; Tillas et al., 2017). Gibson (Gibson, 1977; Gibson, 1979) also believed that affordances were static, that an affordance does not change as the needs and goals of the actor change. As with perception, this aspect of his definition was in opposition to cognitivist views, for whom the value or utility of something is assumed to change as the need of the observer changed (Fayard and Weeks, 2014; Hutchby, 2001; Tillas et al., 2017). This interpretation is limited when applied to technology, because not all affordances are directly perceptible from first use and an element of learning is often needed to successfully use technology (Kaptelinin and Nardi, 2012; Mazé and Redström, 2005; Vyas, Chisalita, and Veer, 2006). This led to a reformulation of the concept within the computing domain when introduced by Norman.

3.2 Norman's Affordances

Affordances were introduced to design and Human-Computer Interaction (HCI) by Norman (Norman, 1988; Norman, 1999; Norman, 2002). For Norman, affordances are properties found in the real world that allow for actions (including the

'climbability' of stairs and clues as to how to use push/pull doors) that can be designed into technological artefacts to provide signals to a user about the purpose and operation of the artefact, e.g., the shading and design of a scrollbar or a clickable button (Kaptelinin and Nardi, 2012; Leonardi, 2011; Norman, 1988; Norman, 2002; Vyas, Chisalita, and Dix, 2017). The position offered by Norman is in contrast to Gibson's view of affordances, as shown in table 3.1. Affordances are processed in the brain rather than directly picked up; affordances are reliant on knowledge, culture and experience, rather than existing independently; and affordances can be learnt whereas for Gibson they cannot (Bardram and Houben, 2018; Kannengiesser and Gero, 2012; McGrenere and Ho, 2000; Oliver, 2005; Torenvliet, 2003; Vyas, Chisalita, and Dix, 2017).

Gaver (Gaver, 1991; Gaver, 1992) also contributed to the development of affordances within computing, building upon Norman's work. Gaver (1991) focused on 'technological affordances' theorising that affordances can be visual but can also be embedded in other senses, e.g. auditory or tactile through sound or vibration; that there are varying types of affordances (e.g., sequential); and that a user's perceptions are embedded in their social and cultural context, personal experience, and intentions: "a cat-door affords passage to a cat but not to me, while a doorway may afford passage to me but not somebody taller" (Gaver, 1991, p.80).

Gaver (1991) noted that the cognitive approaches detailed in computing by himself and Norman are the best for complex and difficult to use artefacts in comparison to Gibson's, that are best reserved for easily-used artefacts. With regards to this research, these approaches are more relevant than Gibson's original affordances, as they pertain to technology use and recognise the impact of contextual factors. However as with Gibsonian affordances, the interpretations of affordances within computing have attracted criticism, particularly in regards to a heavy focus on usability (Still and Dark, 2013; Torenvliet, 2003; Uther and Banks, 2016). When designing trustworthy systems, "the role of systems designers and researchers is thus not one of solely increasing the functionality and usability of

Gibson's Affordances	**Norman's Affordances**
Direct pick up of cues	Cognitive inputs recognised e.g., knowledge, memory retrieval or drawing inferences
Static	Dynamic, changing with context and the needs and goals of the actor
Cannot be learnt	Can be learnt

Table 3.1: Key differences between the ecological psychology approach to affordances by Gibson and the position of Norman when introducing affordances to computing.

the systems" (Riegelsberger, Sasse, and McCarthy, 2005, p.383). Affordances as defined by Norman and Gaver are primarily concerned with usability and ease-of-use (McGrenere and Ho, 2000; Vyas, Chisalita, and Dix, 2017). Whilst usability is of paramount importance to users — there is no possibility of beginning to trust a system if one cannot work it — it is not the sole concern when designing for trust. Additionally, affordances as found within computing have been criticised as they often assume a passive reception of affordances by users (Evans et al., 2016; Leonardi, 2011; Vyas, Chisalita, and Dix, 2017), placing the focus upon designers rather than users when it comes to the interpretation and identification of affordances (Leonardi, 2011; Sengers and Gaver, 2006; Vyas, Chisalita, and Veer, 2006).

Thus far within the theoretical portion of this thesis (figure 1.3), a case has been made for the need to move beyond models of trust in order to inform the design of socio-technical systems. This necessitates an intellectual contribution in the form of translation from the existing concept of affordances to the concept of trust affordances. In the next section, *trust affordances* are theorised, and relevant conceptualisations of affordances post-Gibson and post-Norman are introduced in order to support this development and illustrate how they may be designed into a system.

3.3 Developing Trust Affordances

This research seeks to inform the design of a collaborative research environment (Virtual Lab) whilst embracing the nuance and complexity of trust and accounting for the multiple trustors, trustees, assessments of trustworthiness, and development of trust. In the context of this research, well-placed trust is the primary goal, with trustworthiness cues a prerequisite to achieving this (Liao and Sundar, 2022). Trust affordances seek to contribute to this area, providing insight into how we may design such systems whilst remaining cognizant of the multiple ways of, and reasons for, placing trust. Trust affordances (Thornton, Knowles, and Blair, 2021, p.67)[1] are defined as:

> ***characteristics of the technology by virtue of itself or of features designed into the technology to promote trust by providing access to evidence of (dis)trustworthiness specific to a user, a technology, and their context.***

For the purpose of this thesis, trust affordances maintain focus on informing system design, taking into account the multiplicity of trusts by multiple stakeholders across different dimensions of trust: in the Virtual Lab, in other users, and in data and models. Utilising the definition by Evans et al. (2016), an affordance is:

- not solely affiliated with a specific platform or a feature of a platform;
- variable, dependent on the perception of the user;
- not an outcome, e.g., anonymity, persistence, searchability, and visibility are affordances but privacy is not as it is an outcome.

[1]The phrase 'trust affordances' has been used in literature (cf. De Roure, Goble, Aleksejevs, et al. (2010), Fulmer and Dirks (2018), Kobayashi et al. (2015), and Shin, Zhong, and Biocca (2020)) but had not been formally defined until Thornton, Knowles, and Blair (2021).

3.3.1 Relevant Conceptualisations of Affordances

There are relevant conceptualisations of affordances within the social computing literature that have been instrumental in developing the theory of trust affordances. These concepts deserve attention as they help to alleviate the problems arising from both Gibson's and Norman's approaches above and relate closely to the formation of trust for trustors. These conceptualisations are: social, socio-cultural (collaborative and organisational), and multi-sided affordances.

Social Affordances: Where object affordance refers to properties that determine possible actions given their capabilities, a *social affordance* is the relationship between the properties of an object and a given social group that enable particular kinds of interaction (Bradner, 2001). This conceptualisation was the first to consider affordances beyond an object and a single user. Bradner, Kellogg, and Erickson (1999) illustrate social affordances wherein different users ascribe different meanings to an affordance. For instance, for remote users an affordance opening up communication was perceived as a method to be delegated additional work and therefore the system was unused, whilst for other groups who were co-located with greater social ties, the same affordance was seen as a positive, providing an opportunity to interact with one another (Bradner, Kellogg, and Erickson, 1999). The same system and technological feature had the same affordance, but this perception of the opportunity provided was differentiated by the characteristics of the group. This demonstrates the impact of group composition and context upon the meanings and the usage of a system, and what particular features afford users. As discussed previously, this research adopts a social-psychological approach to trust, recognising that given the same evidence, different users reach different judgements of trustworthiness and different processes when placing trust.

Socio-cultural Affordances: Following the introduction of social affordances, socio-cultural affordances have been also been developed. Bardram and Houben (2018) created *collaborative affordances* including not only the social but also the cultural context of actors that directly influences how affordances are perceived

and how certain interactions and properties afforded the ability to collaborate. Similarly, Vyas, Chisalita, and Dix (2017) extended the concept to a socio-cultural context to create *organisational affordances*. Four conditions are taken as representing context and have a combined influence on the emergence of affordances: technological (functionality and design features); cultural (the beliefs and values held); power (referring to control or law); and interpretative (relating to the knowledge and attitudes of users and the meaning attached to a technological artefact) (Vyas, Chisalita, and Dix, 2017). Vyas, Chisalita, and Dix (2017) note that the convergence of cultural, power, interpretive and technological conditions can lead to the emergence of an unwanted affordance, and an avoidance of using the system. As discussed previously, an understanding of context is fundamental to the formation of trust (section 2.1.2), thus this must also be considered when designing a system with trust affordances. Scientific (sub-)disciplines, as forms of epistemic communities (Knorr Cetina, 1999; Van House, 2002), are likely replete with cultural norms and forms of knowing that shape both the perceptions of affordances and trustworthiness.

Multi-sided Affordances: Islind et al. (2019) coined the notion of two-sided and multi-sided affordances to describe the complexity that occurs when heterogeneous actors co-exist and where affordances can differ for different actors. They argue that whilst collaborative affordances (as discussed above) are affordances shared between both actors, two-sided affordances occur where the *affordances differ between the actors*. In their empirical study, the perception of affordances introduced by virtual consultations was different for nurses and patients, affording visibility and accessibility for the former, and privacy, comfortability, and intimacy for the latter (Islind et al., 2019). Thus affordances of a system not only provide positive actions and social exchanges, but negative experiences too. As illustrated earlier, there are potentially competing interests when it comes to fostering trust between stakeholders and thus there may be two- or multi-sided affordances of a technology. This is explored further in the next section (3.4), wherein the mindfulness and balance

required for multiple stakeholder design is illustrated through the introduction of themes.

In sum, these three conceptualisations of affordances have been instrumental in developing trust affordances, as they illustrate how affordances of systems can work in practice. These affordances reach far beyond object affordances to include social interactions, the influence of context, and the possibility for multiple positive and negative consequences. Importantly, these conceptualisations reveal affordances as relational, emergent, and dynamic. Affordances emerge from the interaction between a user and a technology relative to their contextual setting (Davis and Chouinard, 2016; Evans et al., 2016; Hartson, 2003), leading to multiple possible interpretations of a system (Faraj and Azad, 2012; Hutchby, 2001; Leonardi, 2011). As such, affordances do not operate uniformly and are specific to a user and their context (Davis and Chouinard, 2016; Hutchby, 2001) and may be variable (Anderson and Robey, 2017; Davis and Chouinard, 2016; Evans et al., 2016; McGrenere and Ho, 2000). Bearing a resemblance to the dynamism of trust that ebbs and flows over time and between situations, affordances differ when capabilities change or as new information or skills are learnt (Vyas, Chisalita, and Dix, 2017).

These conceptualisations of affordances have also been important in overcoming the drawbacks of trust models (section 2.6). The static nature of trust models and the need for inclusion of the perceptions of multiple stakeholders was highlighted as a key concern to address. The three conceptualisations of affordances illustrate that we, as system designers, can integrate fluid and idiosyncratic concepts into the design process. The next section continues this trajectory by developing categories of trust affordances from the literature.

3.4 Categorising Trust Affordances

A natural starting point in the design of a technological system for data curation is to focus on intended system users (Lin et al., 2020; Prieto, 2009). Yet this

is challenging when there is no singular user community, as is the case within transdisciplinary research (Cragin and Shankar, 2006; Faniel and Jacobsen, 2010; Grudin, 1994a). As discussed in chapter 2, "trust is the most fundamental but perhaps least well understood property of digital repositories that hold and preserve archival documents" (Donaldson and Conway, 2015, p.2427). Within a Virtual Lab, trust is placed in many and varied ways by trustors, developing "trust from multiple dimensions of associated entities rather than relying on one entity" (Yoon, 2017, p.9). Above, relevant conceptualisations of affordances were introduced in order to account for these multiplicities and to develop *trust affordances*. Within this section, the concept of trust affordances is developed with respect to three categories: transparency, provenance, and social translucency, laying the foundation for further research and problematising current 'solutions' to issues of trust. These themes were derived from the social computing literature with respect to their stated prevalence for contributing to the development of trust (Thornton, Knowles, and Blair, 2021). The purpose of this section is not to create specific trust affordances, which will be derived from the empirical data, but rather to provide insight in order to understand the nuance, especially given the common assumptions made within the wider literature.

3.4.1 Provenance

Provenance relates to the lineage of data, from the data producer through aggregation with other sources or transformation of variables to data curation and eventually ending up with the data consumer (Bates, Lin, and Goodale, 2016; Cheney, Chiticariu, and Tan, 2009; de Lusignan et al., 2011; Van House, 2002). Documentation in the form of meta-data is commonly thought to satisfy this need for provenance by providing a historical record (de Lusignan et al., 2011; O'Hara, 2019; Zhao et al., 2011). From a single data source and data producer, provenance information may be useful in determining the trustworthiness of a data producer, and extrapolating this trust to the data they produce (Gamble and Goble, 2011;

Knowles, Rouncefield, et al., 2015). However, complexity quickly compounds when we consider that a dataset may be a combination of data sources by different data producers and curators for different purposes (Barclay, Preece, Taylor, and Verma, 2019; Bowker, 2000; Faniel, Frank, and Yakel, 2019; Marsh, Atele-Williams, et al., 2020). Provenance information is thus either impossible to create, when there is no clear path from data producer through data curation to the data consumer, or likely replicates the complexity within the information provided.

It is also the case that provenance is not necessarily a marker of trustworthiness if it not verifiable or reliable in itself (Knowles, Harding, et al., 2014; Lyle and Martin, 2010), e.g., by allowing the data consumer to validate for themselves (Marsh, Atele-Williams, et al., 2020; Sexton et al., 2005). Within environmental data science, this opportunity to scrutinise data, models, and methods has been recognised as a necessary requirement for systems (Henrys and Jarvis, 2019; Meyer, Weigelt, and Kreft, 2016; Pescott, Humphrey, and Walker, 2018). In this regard, meta-data alone cannot solve all provenance queries to all users who may utilise different strategies to form trust (Birnholtz and Bietz, 2003; Donaldson and Conway, 2015; Marsh and Dibben, 2003). Thus in addition, provenance may be evidenced through a trusted third party, i.e., a technological system (Barclay, Preece, Taylor, and Verma, 2019; Smith, 2000), such as a repository or a Virtual Lab. These data infrastructures can facilitate access to data and importantly provide guarantees of authenticity and quality either by the documentation they proffer, e.g., with standards that require information to be provided upon data upload (Hollaway et al., 2020; Ross and Mchugh, 2006; Yakel et al., 2013) or that data and associated records will be preserved and archived (Borgman, 2015; Clarke et al., 2006; Jantz and Giarlo, 2007; Lin et al., 2020).

Provenance is thought to imply authenticity and integrity, proving that the data is the same as it was upon collection (Lynch, 2000; Sexton et al., 2005), and is thus trustworthy as a result of this (Penner and Klahr, 1996; Zhao et al., 2011). Data quality is signalled as a prevailing problem in data science (Gamble and Goble,

2011; Reis et al., 2015; Van House, Butler, and Schiff, 1998), and provenance information is thought to aid interpretation and assess quality (Borgman, 2015; Bowker and Star, 2000; O'Hara, 2019; Zuiderwijk, 2017). The underlying belief is that provenance is thought to provide context regarding the source of the data and a record of the assumptions and decisions made (Faniel, Frank, and Yakel, 2019; Pescott, Humphrey, and Walker, 2018; Van House, 2002), thus enabling repeatable and reproducible science using the data (Bechhofer et al., 2013; de Lusignan et al., 2011; De Roure, Goble, Aleksejevs, et al., 2010) and hence, trust (Lord et al., 2004; Lyle and Martin, 2010; Zhao et al., 2011). Yet, context is not necessarily a given. Preservation relies upon *usable* and *meaningful* curation, rather than simply preservation (Rothenberg, 2000). There are different types of provenance that afford trust in different ways. Evidential or 'coarse-grained' provenance provides information on ownership and history, whereas informational or 'fine-grained' provenance provides contextual information surrounding the data (Cheney, Chiticariu, and Tan, 2009; de Lusignan et al., 2011; Hertzum, 1999; Spiekermann et al., 2019). The differences between these two forms indicate separate values. Evidential value relates to structure, procedures, and proving ownership, i.e., *where* has data come from; and informational value refers to the contents of the document for reference, contemplation and research, i.e., the *why* including the underlying assumptions and purpose for which it was created (Cheney, Chiticariu, and Tan, 2009; Hertzum, 1999). Within the literature, assumptions are made by simply invoking 'provenance' without further specification in its operationalisation and a simple correlative belief that information equates to trust.

Under a trust affordances perspective it is unclear what type of provenance is needed, for whom, and for what purpose. For a data producer, providing provenance information for those in their community with the same shared (and thus often unexplained) terminologies and assumptions and the amount of context required (Bowker, 2000; Brown and Duguid, 2001; Grudin, 1988; Markus, 2001) would likely be different to those beyond this community, and may equate to

different forms of provenance, e.g., by taking common disciplinary conventions and meanings for granted and not explaining them in full detail. Documentation affords the preservation of information and transfer beyond its origin (Hertzum, 1999; Piorkowski et al., 2021; Sundqvist, 2011) and a static, immutable snapshot of the data at a certain point in time (Arnold et al., 2019; Bates, Lin, and Goodale, 2016; Gebru et al., 2021) but this does not necessarily equate to trust for different users: to whom is the provenance information directed? Are they trained in the same discipline and familiar with the data and tools? What do they need to see to determine provenance? How does it need to be presented? Do they need further context? Is the provenance information provided understandable and meaningful?

3.4.2 Transparency

Transparency connotes visibility and openness, i.e., how a model works, how a system functions, or how a decision is made, and is linked to intelligibility, explainability, and understanding (Arya et al., 2019; Ashoori and Weisz, 2019; Preece et al., 2018; Vaughan and Wallach, 2021). Transparency is linked to trust in repositories (O'Hara, 2019; Ross and Mchugh, 2006; Yakel et al., 2013) and trust in data (Faniel, Kriesberg, and Yakel, 2016; Sexton et al., 2005). This connection rests upon the assumption that transparency enables us to determine trustworthiness, as we cannot trust that which is hidden. One of the main mechanisms for achieving transparency is documentation (Gebru et al., 2021; Hind et al., 2020; Holland et al., 2018). The underlying assumption rests upon the belief that knowledge equates to trust. Yet, an absence of useful information can impair perceptions of trustworthiness, and equally, too much information does not foster trust (Harper, 2000; O'Neill, 2020; Pieters, 2011). Rather, an abstraction of information is required (Castelfranchi and Falcone, 2020; Kiani et al., 2019; Markus, 2001; Star, Bowker, and Neumann, 2003). This is referred to as 'meaningful transparency' (Norval et al., 2022; Obar, 2020) or 'layered transparency' (Sexton et al., 2005) referring to a selectivity by data producers or curators on behalf of the data consumer to tailor

information in different ways, and thus presenting the same information in different layers and aiding understanding and accessibility in ways that are meaningful to users such that they can derive insight relevant to their needs (O'Neill, 2018; Preece et al., 2018; Rothenberg, 2000). In their creation of a virtual research environment for environmental data science, Elkhatib, Gemmell, et al. (2019) provide access to different levels of information for varying user groups: scientists, policy makers, farmers, and the general public. Whilst this may afford trust for data consumers, a tension may arise in conflict with data producers. This approach assumes that they have both the knowledge of biases, potential consequences, and application to different contexts as well as the time and capability to undertake this translation work (Poursabzi-Sangdeh et al., 2021; Raji et al., 2020; Richards et al., 2021; Sokol and Flach, 2020). This may result in a tension between what data consumers require available to them in a Virtual Lab, and what data producers are willing to produce to aid in data reuse.

Invoking a multi-sided affordances view, it is also the case that for data producers, transparency may lead to a diminished perception of trustworthiness in a system. There is a fundamental tension between transparency and privacy (Ackerman, 2000; Bannon and Schmidt, 1989; Rodden and Blair, 1991). Too much visibility may lead to lack of disclosure and sharing for a number of reasons (Cook and Santana, 2020; Schiff, Van House, and Butler, 1997; Suchman, 2002). It may be confidential or under embargo (Birnholtz and Bietz, 2003; Markus, 2001), they may be concerned with being credited (Cragin and Shankar, 2006; Hine, 2006; Stall et al., 2019; Zimmerman, 2008a), being able to publish their own results first (Wallis, Rolando, and Borgman, 2013), or that their data may not be used for the right context (Birnholtz and Bietz, 2003; Faniel and Zimmerman, 2011). Thus whilst it is clear that transparency can aid in fostering trust for some, for others this ideal does not foster trust. Even within their own collaborations, researchers want the autonomy to decide with whom and when they share their research (Schmidt, 2000). What form this should take within a Virtual Lab also remains unclear: how

much transparency is 'enough' for trust? Are there different acceptable levels of transparency for different stakeholders? How does transparency lead to (loss of) trust? Who is responsible for developing different levels of information?

3.4.3 Social Translucency

Many in science work in isolation and shared spaces can transcend the distance between them whilst also including other stakeholders, such as policy makers and members of the public (Elkhatib, Blair, and Surajbali, 2013; Zuiderwijk, 2017). Yet there are social barriers to data sharing and data use in which trust may be needed (Birnholtz and Bietz, 2003; Faniel and Jacobsen, 2010; Van House, Butler, and Schiff, 1998). Trust is theorised to aid in data sharing by increasing the levels of disclosure from one to another, and also reducing the screening of knowledge from others (McEvily, Perrone, and Zaheer, 2003). Many of the problems addressed above with respect to provenance and transparency also relate to the social side of scientific practice. For instance, regarding the transfer of context, documents afford the transfer of data in time and space, but do not allow us the opportunity to ask questions and receive answers (Chin and Lansing, 2004; Hertzum et al., 2002; Levy, 2000; Olson and Olson, 2000). An exchange of words can often help to foster trust (O'Neill, 2002) but only if a system is designed such that it affords communication. In a face-to-face conversation if we can see a puzzled look upon the person with whom we are talking, we can tailor the content of our communication accordingly to align with their knowledge or experience. Yet in online environments, these social cues are absent (Erickson and Kellogg, 2000; Luhmann, 2017; Myskja, 2008; Riegelsberger, 2002) and the trust signals we use in real-life may be different or unavailable (Riegelsberger and Sasse, 2010). In computer systems, awareness and visibility of others' actions are designed into the technology to combat these issues (Dourish, 2001; Finholt and Olson, 1997). Erickson and Kellogg (2000) call a system 'socially translucent' if it has three characteristics: visibility, awareness, and accountability. Whilst Erickson and Kellogg (2000) do not

directly connect social translucence to trust, there are many crossovers. For instance, within transdisciplinary scientific research, collaboration beyond disciplinary silos is essential (Blair, Leeson, et al., 2019; Edwards et al., 2011; Neff et al., 2017), and inter-disciplinary dialogue is a necessity (Beagrie, 2006; Blair, Beven, et al., 2019; Jeffery, Candela, and Glaves, 2020). Proximity affords opportunities for collaboration whereas distance can prevent it (Bradner and Mark, 2002; Hollan and Stornetta, 1992). Virtual research environments can be designed to 'augment' and 'relax' these obstacles (Finholt and Olson, 1997; Olson and Olson, 2000) through rich and varied modes of communication (Erickson and Kellogg, 2000; Hollan and Stornetta, 1992).

Yet, as with the other affordance themes discussed above, there are challenges to this. Transferring tacit knowledge is complex (Collins, 1992; Markus, 2001; Sellen and Harper, 2001). Whilst forms of documentation are proposed to enable the transfer of knowledge across time and space, both Edwards et al. (2011) and Faniel and Jacobsen (2010) found that direct contact and follow-up questions were preferred and desired by data consumers. Social exchanges are difficult to scale to a system-level, and may face resistance from data producers as the number of conversations they are drawn into may be more than they wish (Faniel and Zimmerman, 2011; Wallis, Borgman, et al., 2007).

Trust plays a fundamental role in facilitating and continuing collaboration (Costa, Fulmer, and Anderson, 2018; Marsh, Atele-Williams, et al., 2020; Newell and Swan, 2000). However, the majority of work in this area assumes that — whilst geographically or temporally distant — the collaborators are known. Identity is frequently assumed to be fundamental in the assessment of trustworthiness as it provides the trustor with an ability to frame previous experience and history (Lewicki and Bunker, 1996; Nissenbaum, 2001; Turilli, Vaccaro, and Taddeo, 2010). Further to this, identity also enables us to ascertain the reputation of the trustee (Artz and Gil, 2007; Talboom and Pierson, 2013). A particular data producer may be more reliable than another (Jirotka et al., 2005; Zimmerman, 2008a) or we might

see that someone we trust has used — and therefore endorsed — a dataset (Ehsan et al., 2021). Designing systems for known groups of users *that also know each other* introduces limits and boundaries that designers can work with and assumptions to be made (Grudin, 1994a; O'Day and Nardi, 2003; Schmidt and Bannon, 1992; Van House, Butler, and Schiff, 1998). Such a system will require different features — and produces the opportunity for the perception of different affordances — than a system in which there are unknown users with unknown familiarity and relationships. In lieu of these certainties, assumptions are harder to pose: how can we determine the level of familiarity for multiple stakeholders? How does trust aid in collaboration currently and (how) will a Virtual Lab alter this? Do the social translucent system characteristics relate to this research context? What forms of communication will work for stakeholders within a Virtual Lab? How can we meaningfully represent identity, visibility and awareness? Will stakeholders' preferences conflict?

3.5 Chapter Takeaways

Virtual Labs are differentiated from data repositories, which typically focus on meta-data, provenance, classification and standardisation (Borgman, 2015). They incorporate the traditional features of repositories, alongside tools for conducting experiments and analysis, generating visualisations, or including elements of social networks (Bechhofer et al., 2013; Elkhatib, Blair, and Surajbali, 2013; Hallgren et al., 2016; Zhao et al., 2011). Beyond this, Virtual Labs enable the abstraction of information tailored to skill levels (Gordov et al., 2018; Hollaway et al., 2020) and the inclusion of glossaries and background information (Hallgren et al., 2016). By invoking a trust affordances perspective, i.e., a deconstruction of assumed or implied features to foster trust, a technology can neither be said to afford trust in and of itself, nor can it's features be taken for granted in affording trust for all users. Clearly, and as argued throughout this research, managing this complexity requires nuance.

Affordances can be utilised to *guide the design of technology* capturing the relationships and interactions between users and an artefact with an emphasis on the context that these interactions are situated in (Bradner, 2001; Gaver, 1991; Hartson, 2003; Kannengiesser and Gero, 2012). By expanding this concept further to trust affordances, this research contends that designers can invoke this approach to develop a system that is trustworthy-by-design (Knowles, Harding, et al., 2014) *whilst* bringing the multiplicity and dynamics of trust for different stakeholders into focus, as demonstrated by the research aims and questions (chapter 1). Importantly, trust affordances enable a deeper exploration of each component of trust models, providing us much needed insight into each 'box' (Castelfranchi and Falcone, 2010). This chapter signals the major intellectual contribution of this research: developing trust affordances and the related thought processes to an area in which there are limited approaches, design principles, or design frameworks to follow or models of trust to utilise. This provides a foundation to address the overall goal of this research (figure 1.2) which will be developed through the empirical portion of this work.

Chapter 4

Research Approach

Key to addressing the research questions, aims, and objectives is the adoption of a suitable research design. As shown in figure 1.3, this chapter provides the outline and reasoning for the the methodological undertaking of the research including justification for the decisions made, which is then undertaken within empirical chapters 5 and 6. As shown in figure 4.1, this chapter firstly details the case study methodology adopted (section 4.1) and expands upon these cases (section 4.2) before outlining the qualitative methods for data collection and analysis used (section 4.3). This chapter will provide an overview of the cases selected, the data collection and analysis routines performed, and the suitability of the research design to answer the research questions and aims.

4.1 Research Design

This research adopts a case study approach using a constructivist philosophical perspective. Case studies facilitate an in-depth exploration and examination of complex phenomena within a specific context (Rashid et al., 2019; Saunders, Lewis, and Thornhill, 2000). Case studies are useful when the research questions concern 'how' and 'why' (Stake, 1995; Yin, 2017), with an in interest in particulars and complexities rather than generalities e.g., decisions made, and the contextual factors

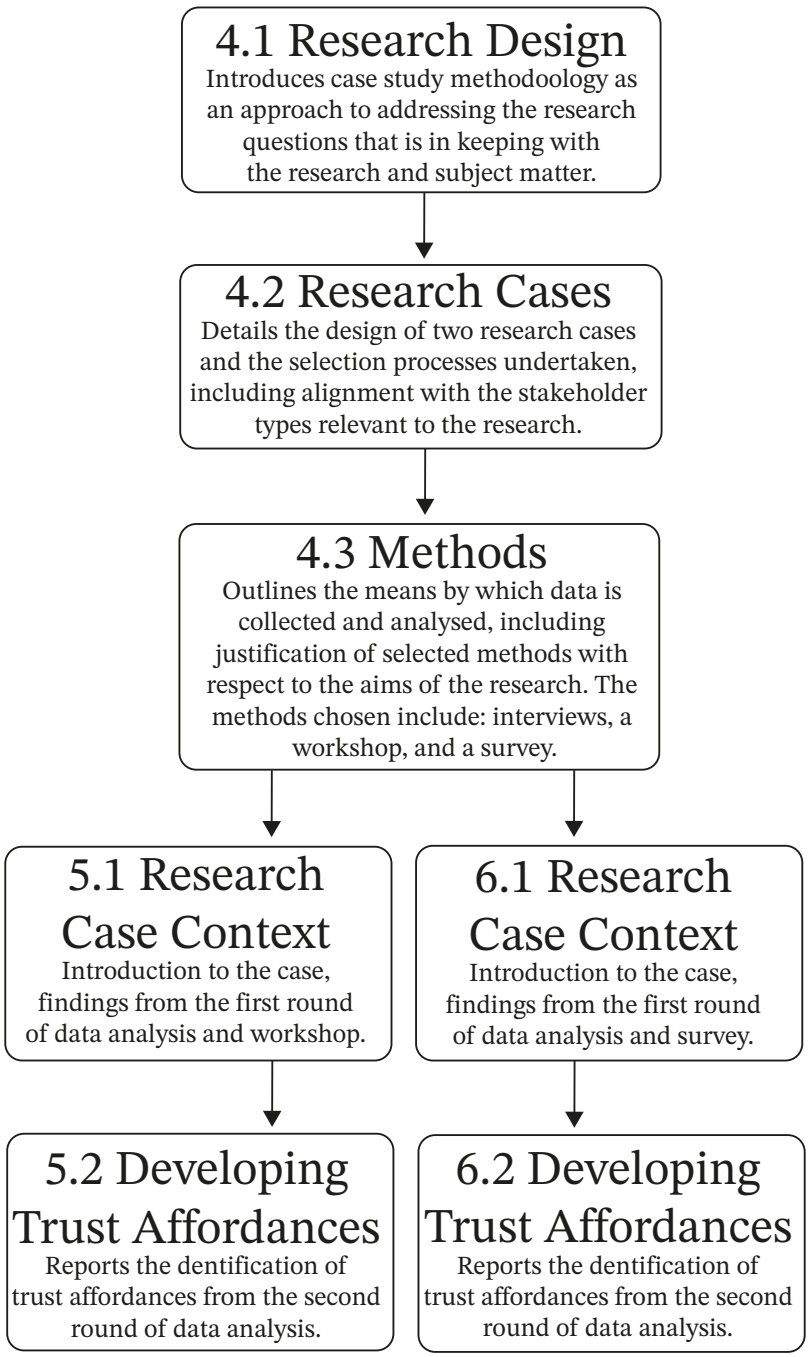


Figure 4.1: Structure of the methodological and empirical chapters.

that influence decision-making (Baxter and Jack, 2008; Denscombe, 1998). This approach has been utilised for research aims similar to the research at hand. Carlson and Anderson (2007) utilise a case study approach to investigate the re-use of data

for four different e-science funded research projects noting their ability to develop richly illustrated themes, and similarly, Winters and Mor (2009) note that the use of case studies were useful in the creation of design patterns, as they sit halfway between approaches that result in too much context-specific data and too much abstraction.

Case studies are compatible with numerous philosophical perspectives and methods (Baxter and Jack, 2008; Boblin et al., 2013; Creswell, 2013), including the paradigm adopted within this research, constructivism (Kivunja and Kuyini, 2017; Stake, 1995). A case study approach advocates for the use of a variety of data sources, data types, and data analysis methods (Boblin et al., 2013; Denscombe, 1998; Petty, Thomson, and Stew, 2012) to create a view of reality from different angles, i.e., multiple interpretations and explanations of one phenomena (Davis, 2008; Tracy, 2010). The aim of constructivist inquiry is to understand, interpret, and analyse the social world of participants within a specific context (Bryman, 2008; Leeming, 2018; Mertens, 2014). A paradigm characterises a researcher's worldview and beliefs that guide their research (Guba and Lincoln, 1994; Saunders, Lewis, and Thornhill, 2000). The epistemological, ontological, and axiological positions of a constructivist paradigm are detailed in Appendix A.1. These positions reflect the aims of this research: to understand the perceptions of participants (epistemology) and the multiple meanings attributed to a phenomenon (ontology). An overview of the corresponding relations between the social-psychological approach to trust, constructivist paradigm, and qualitative methods are detailed in Appendix A.2, illustrating alignment with a qualitative case study approach to study trust within environmental data science (Hyett, Kenny, and Dickson-Swift, 2014; Rashid et al., 2019; Priya, 2021). This approach is consistent with CSCW (Computer Supported Cooperative Work), which has an epistemological concern for understanding users and their contexts (Button, 2006; Crabtree, Nichols, et al., 2000; Erickson, 1996; Suchman, 1987).

4.2 Research Cases

The distributed nature of the groups of potential users can make identifying a representative range of users challenging (Schiff, Van House, and Butler, 1997). Stakeholders in environmental data science include: *data producers* who create data, *data consumers* who use data, and *data curators* who manage data. Multiple case study design was adopted to explore the richness, depth and complexity about one phenomenon of interest from multiple angles and perspectives (Bryman, 2008; Lauckner, Paterson, and Krupa, 2012; Yin, 2017). To represent different trustors and trustees data science pipeline (figure 4.2), individuals within specific communities were chosen across two case studies (Denscombe, 1998; Lauckner, Paterson, and Krupa, 2012; Stake, 2005).

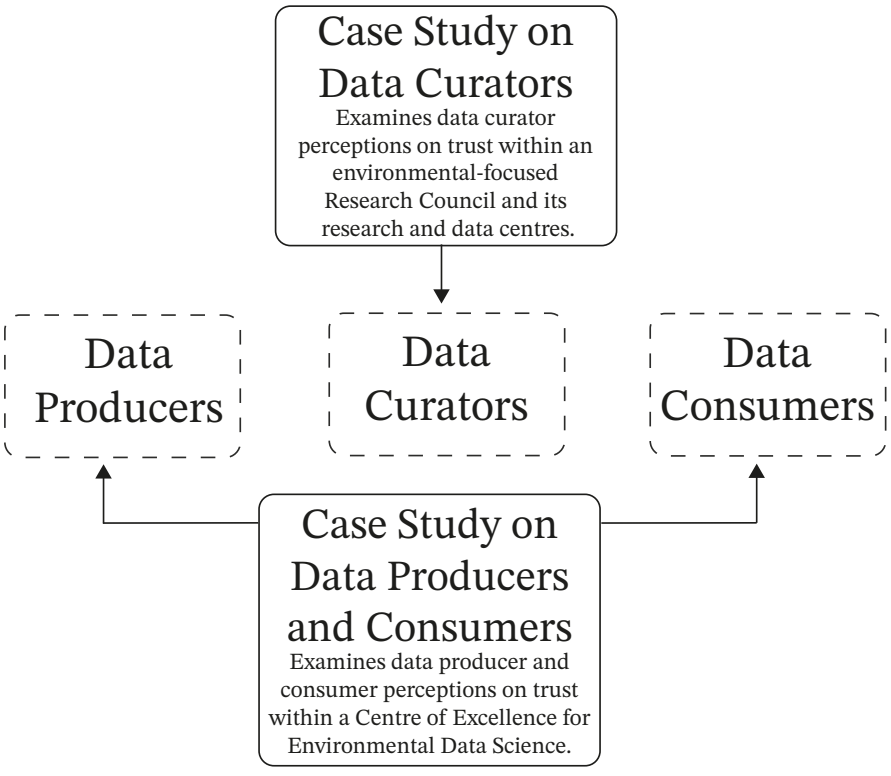


Figure 4.2: How the research cases relate to the different types of stakeholders: data producers, data curators, and data consumers.

The two cases chosen are complementary, focusing on different perspectives of stakeholders within environmental data science and thus able to address the research questions from a variety of perspectives. The first case focuses on data curators, and is designed to capture the perceptions of senior representatives from research centres and data centres within NERC (Natural Environment Research Council). As a research council, NERC has six research centres covering the wide array of environmental science disciplines within its remit: British Antarctic Survey; British Geological Survey; National Centre for Atmospheric Science; National Centre for Earth Observation; National Oceanography Centre; and UK Centre for Ecology & Hydrology. Connected to these research centres are data centres. The data centres focus on different environmental disciplines: marine (British Oceanographic Data Centre); atmospheric, earth observation, and solar and space physics (Centre for Environmental Data Analysis); terrestrial and freshwater (Environmental Information Data Centre); geoscience (National Geoscience Data Centre); polar and cryosphere (UK Polar Data Centre). Four participants were selected for their role as data curators with whom project members had connections, with each participant having past and/or current experience as data producers and consumers. These participants were managers or leads in NERC data centres, with some participants also having leading roles within the affiliated research centres. The four participants had similar self-described roles: the head of a NERC environmental data centre; a curation manager for a data centre; the head of a data centre and the head of a team within the related research centre; and as the lead of a team of data managers. A fifth participant was a representative for NERC as a whole, rather than aligned with a specific research centre. The final participant was not a data curator, but had significant experience as a producer and consumer within environmental data science and a strong connection to one of the research centres and the associated data centre. Participants are illustrated in figure 5.1. The second case focuses on data producers and data consumers. Those selected for this case are experienced environmental data scientists within a Centre of Excellence. Within the

Centre, theme leaders are arranged according to a matrix of environmental themes (air, biodiversity, land and soil, and water and ice) and methods (data acquisition and infrastructure, data science methods, decision-making under uncertainty, and complex systems). These theme leaders were chosen as potential participants, given their esteem to be chosen to fill positions within the Centre of Excellence. Seven were recruited in the former and six were recruited in the latter group (fifteen were invited in total, but two either declined or did not respond). This is illustrated in figure 6.1. Many of those interviewed were from their own specific sub-discipline, and were thus able to provide reflections upon their own experiences within their disciplines and as part of a collective group of environmental data scientists.

A recognised problem within environmental data science are the small and fragmented communities from multiple sub-disciplines (Blair, Bassett, et al., 2021), making the process of delineating a clear case challenging. A small but homogeneous sample was selected for each case using purposeful, theoretical sampling (Crang and Cook, 2007; Roy et al., 2015; Sandelowski, 1996). This approach maximises variation when the sample is small (Bryman, 2008; Creswell, 2013; Patton, 2002; Ritchie and Lewis, 2003) and enables the representative selection of a variety of participants in order to gain a range of perspectives (Creswell, 2013; Mertens, 2014; Vasileiou et al., 2018). The selection of cases were bounded temporally and by group membership, distinguishing those within and outside of the group (Denscombe, 1998; Yin, 2017). The cases had to fit in with the length of the research project and had to have relevance to the research, i.e., be involved in the practice of environmental data science. These two cases were selected to represent each type of trustor and/or trustee within environmental data science and thus directed whom to invite for participation. As illustrated in figure 4.2, the first case focuses primarily on data curators, with the second focusing on those with experiences as data consumers and data producers. For both cases, a wide range of participants were gained, i.e., no two participants were from the same sub-discipline with the same research focus within each case. A cohesive and identifiable group is beneficial because of the

challenge in design when users are distributed and certain assumptions cannot be made (familiarity, expertise, etc.) (Grudin, 1994a; O'Day and Nardi, 2003; Schmidt and Bannon, 1992). By grouping those within a case, a strict boundary can be drawn aiding the conceptualisation of trust for different stakeholders and "between whom and for whom issues of trust arise" (Knowles, Harding, et al., 2014, p.1062).

4.3 Methods

Mixed methods — the use of different qualitative or quantitative methods, or a combination thereof (Bryman, 2008) — were chosen to capture trust from multiple angles. As a subjective phenomena, trust is experienced uniquely and has individual-specific meanings (Clément, 2020; Lewis and Weigert, 2012; Marsh and Dibben, 2003; Sundqvist, 2011), and is often invisible, going unnoticed until trust issues arise (Baier, 1986; Botsman, 2017; D'Cruz, 2020; Rieder and Simon, 2016). Consequently, determining how to research trust can be challenging (Blomqvist, 1997; Jirotka et al., 2005; Lyon, Möllering, and Saunders, 2015; Marsh and Dibben, 2003), particularly in eliciting precise meanings from participants (Button, 2006; Lagerspetz, 2014; Sturdee et al., 2021). Qualitative methods are appropriate for such circumstances, by drawing out these meanings (Creswell, 2013; Law, 2004); gaining an understanding of how trust is experienced (Jones and George, 1998; Myskja, 2008; Simpson, 2012); and capturing the richness and subtlety of voices and positions, revealing nuance and detail (Maxwell, 2009; Sandelowski, 1996).

Interviews are the primary method use in both cases. An additional method is undertaken in both cases, a workshop and a survey in case one and two respectively, in order to assess emerging themes and findings (Priya, 2021). Two doctoral supervisors were also communicated with throughout the process as a form of peer debriefing, when someone who is familiar with the research or the phenomenon being explored is consulted (Creswell and Miller, 2000; Lincoln and Guba, 1985).

Regarding data analysis, and reflecting the flexible and emergent nature of

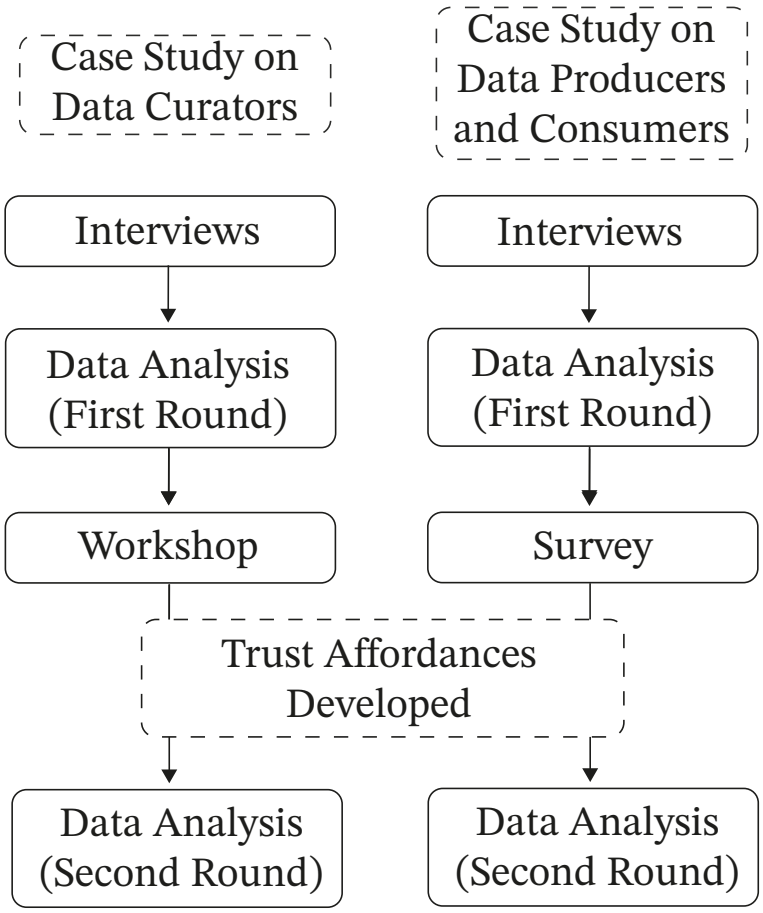


Figure 4.3: Process of data collection and data analysis for both case studies.

qualitative research, the concept of trust affordances was developed following initial data collection and data analysis, thus a second-round of data analysis was included to validate the concept and create specific instances of trust affordances (figure 4.3). The second round of data analysis was conducted when sufficient time had passed to enable exploration of trust affordances. The interviews were the primary data analysed in two rounds for each case study, with the workshop and survey used to evaluate data saturation. The remainder of this section details the data collection and analysis for each method in turn.[1]

[1]Prior to data collection, all research designs were successfully approved by Lancaster University Faculty of Science and Technology Ethics Committee. All participants received an information sheet. Individual participants had to sign individual ethics forms for the interviews, were asked

4.3.1 Interviews

A semi-structured interview approach is taken, enabling the dynamic and unique character of each interview to be formed (Bryman, 2008; Kvale, 1996; Mason, 2002; Rubin and Rubin, 2005). Semi-structured interviews provided 'rich' and 'thick' description, which involves deep, detailed accounts: of settings, participants, and themes (Creswell and Miller, 2000; Denzin, 1989; Lub, 2015; Fusch and Ness, 2015; Geertz, 1973; Roy et al., 2015). Thick descriptions are important to gather the specifities of a particular context (Maher et al., 2018), pertaining to the goals of this research.

Interviews within the Research Councils case were undertaken as part of a research project.[2] My role involved independently planning, conducting, analysing and presenting the interview results and publishing the final work (Thornton, Neumann, et al., 2019). Six semi-structured one-hour interviews were undertaken to identify the challenges that data centres face, and those of the data producers and consumers they represent in order to explore how mechanisms to foster trust can be identified and formalised into distributed ledger technology. Interviewees were either related to the project (n = 2) or were introduced via members of the project who were then emailed for participation (n = 4). All interviews were conducted either by telephone (n = 4) or in person (n = 2) due to geographical proximity and ease for participants. A copy of the participant information sheet, e-mail invite, interview schedule, and consent form can be found in Appendix B.

Thirteen one-hour interviews were undertaken within the Centre of Excellence case to explore trust for environmental data producers and consumers.[3] Those

for consent in the workshop, and were given a consent page for the survey. All data management protocols were adhered to.

[2]EnvChain, 2018, EIDC NERC Data Innovation funding award. `https://tinyurl.com/envchain` accessed June 2023.

[3]These interviews were all conducted virtually between April and July 2020 due to the COVID-19 pandemic. This was undertaken with another doctoral researcher as our topics are related but separate we: submitted our own ethics applications; asked our own questions; divided the

selected to interview had experience as established researchers, i.e., not Early Career researchers (less than seven years since PhD completion). As detailed above, they were chosen as a representative sample across environmental data science, and this case focused on the role of existing data repositories and the potential for Virtual Labs. A copy of the participant information sheet, e-mail invite, consent form, and interview schedule can be found in Appendix C.

The number of participants interviewed is reflective of, and consistent with, qualitative research within the field of data studies: n = 14 earthquake engineers and data reuse (Faniel and Jacobsen, 2010); n = 13 ecologists and data reuse (Zimmerman, 2008a); n = 12 natural resource management agency staff and trust (Sharp et al., 2013); n = 19 researchers with data repository experience and trust (Yoon, 2014); and n = 22 interviews regarding trust in data within scientific digital libraries (Wallis, Borgman, et al., 2007).

All empirical data in the interviews was recorded and transcribed verbatim, either manually with a dictaphone and word processor in the Research Councils case, or using a combination of automation and manual transcription in the Centre of Excellence case.[4] Following the data management plan set out in the ethics application, data was transcribed and then the audio recording deleted, with the remaining typed transcription being anonymised and stored securely on encrypted devices (Creswell, 2013). For clarity in cross-comparison between cases, participants in the Research Councils case were anonymised alphabetically (A-F) and in the Centre of Excellence case, numerically (1-13).

To analyse the data, the data was read through and then openly coded inductively first, as a 'first cycle' or 'first pass' (Braun and Clarke, 2006; Saldaña, 2013; Sharp et al., 2013). The coding process involved systematically labelling concepts, themes, and lower-level terms in order to reduce and make sense of the

transcribing between us; and then conducted our own separate analyses.

[4]In the second case, interviews were conducted using online meeting software: Microsoft Teams, Skype, or Zoom which provide automatic transcription, but these were then manually cleaned again by listening to the recording and altering/correcting the transcription.

data (Creswell, 2013; Rubin and Rubin, 2005). This coding cycle was done manually and from the bottom up (Clark et al., 2007; Crabtree and Miller, 1999; Wilson and Chaddha, 2009), meaning that each transcript was read line-by-line and each meaningful section of the text was assigned a label (code) in a text editor (Rubin and Rubin, 2005). Following this, the codes were reviewed, sorted and reassembled into groups (Creswell, 2013; Rubin and Rubin, 2005). Codes are analysed and considered with respect to higher-level overarching themes (Braun and Clarke, 2006). In the first case, this involved creating new documents in Microsoft Word for each theme, and new groups using coding software NVivo[5] for the second case.

A second cycle of coding was conducted to achieve "a more detailed and nuanced account of one particular theme, or group of themes, within the data" (Braun and Clarke, 2006, p. 83). As illustrated in figure 4.3, this second cycle was conducted following the development of trust affordances. This process of two cycles is illustrative of what Creswell (2013) calls 'layering analysis': from general to particular (or vice versa). This second pass did not involve the application of pre-determined or *a priori* codes, but did involve re-analysing the data in terms of trust affordances, reviewing the data from a different angle to picture the data differently. Interviews are best-suited to developing an understanding of people think and feel about trust (Möllering, 2013). Because context and description are important given the paradigm and methodology, and part of the research aim to gain a formative understanding of the context, this combination was beneficial in helping to achieve the research aims, allowing for the same data to be viewed at different angles and new insights derived.

4.3.2 Workshop

In the Research Councils case, the first cycle of data, i.e., the interviews, were analysed as detailed above and then presented at a workshop for a wider range of stakeholders (n = 17), all of whom were related to NERC and its different research

[5]NVivo for Mac v.1.6

institutes. Those invited to the workshop consisted of NERC staff members and those who worked across the broad range of research centres and data centres within NERC's remit. In terms of data collection, no audio recordings were taken and anonymous notes were typed up manually throughout the day. Registrants for the workshop were aware upon acceptance of invitation and on registration on the day that photos and results of the workshop would be taken, and were given the option to opt out. No names were collected when feedback was captured during the workshop. The workshop was structured so that feedback was attained through sticky notes from each participant and verbally through group feedback to the room. This research was presented in the first session, and workshop participants were asked to think about "the things that have resonated with you" and "the things that you haven't heard". Rather than a thematic analysis as with the interviews, the typed up notes and sticky notes (Appendix E) were reviewed after the workshop and compared to the interview data to evaluate the themes found.

4.3.3 Survey

In the Centre of Excellence case, a survey was utilised to validate the interview findings. The survey consisted of questions using a combination of Likert scales and short-form answers. The Qualtrics platform[6] was used to collect and store data using a Lancaster University account, no participants names were taken. In the survey, participants were given an online information sheet, consent form with tick-box to express consent and to opt in or out, if participants gave their consent they were directed to the survey and if they opted not to consent they were directed to exit the survey. Participants were recruited via a mailing list for the Centre and a link to the survey was provided at the end of a seminar presentation given on the topic. The invite stated that those who had been interviewed should not complete the survey to avoid similar results being collected. Approximately 70-80 people would have

[6]A preview of the questionnaire can be found here: `https://tinyurl.com/qualtrics-lvmt` accessed June 2023.

received the invite to participate, though the survey only received 10 responses. A screenshot of the consent form from the Qualtrics platform and a copy of the survey schedule can be found in Appendix D.

4.4 Chapter Takeaways

A dynamic and detailed research design was chosen to match the nature of trust. As a concept, trust is so unique and yet universal, that qualitative methods are undoubtedly invaluable as an empirical tool. A great deal of methodological research was undertaken in order to design an approach that can effectively address the aims of the research, and to ensure that the adaptation of traditional methods, in particular the inclusion of second analytical rounds, would be suitable. The approach to trust and philosophical paradigm do not dictate a choice of methods, and there were other methods planned and ethical approval gained for this research. This included both ethnographic fieldwork and a prototypes study forming separate empirical data collection and analysis, however due to the COVID-19 pandemic these were cancelled. Whilst these methods may have contributed to research legitimacy and credibility, future work using this research has already begun (described in further detail in chapter 8).

Research is flexible and changes throughout its course. This research started out by exploring trust in data, basing this exploration upon environmental science. However, it became apparent that to focus on one dimension of trust would lead to the neglect of other aspects, e.g., trust in people. This research instead takes a broader focus, trust in data includes trust in people, trust in systems, and trust in institutions. The design of technology to support trust in data necessarily involves these other elements and *should be considered* within design. Taking the time to re-scope and reflect has positively influenced the outcomes of this research. Similarly, trust affordances were not considered at the outset of the research, but were borne out following an examination of theory (trust models) and the empirical

data collected. Thus, it is important to recognise the journey of research, as detailed in figure 4.3, and how this contributes to the overall research outcomes.

A suitable and appropriate research design is paramount to addressing research questions. A considerable amount of time with a great deal of care and attention has been spent crafting the research approach detailed within this chapter. This has resulted in a commensurate and logical flow between the approach to trust, research paradigm, and research design providing the foundation needed to answer the research questions and overall aim of the research. The following two chapters present the findings and analysis (figure 1.3) of the two research cases as outlined within this chapter.

Chapter 5

Exploring Data Curator Perceptions: Trust within a Research Council and its Research and Data Centres

The Natural Environment Research Council (NERC) funds and supports environmental research within the UK, with a key focus on the preservation and provision of quality data to facilitate this work. This includes six research centres and five data centres (section 4.2). Data centres were previously funded separately, but in April 2018 these were integrated into a single service, called the Environmental Data Service.[1] Research centres are still connected to and guide the respective data centres previously under their sole remit.[2] The aim of this transformation is to increase access and efficiency, enhance curation, and improve cross-disciplinary collaboration to improve the process of conducting science.

The research reported within this chapter builds upon a funded feasibility project exploring the potential application of distributed ledger technology (blockchain) by

[1] `https://eidc.ac.uk/policies/digitalPreservation` accessed February 2024.

[2] `https://eds.ukri.org` accessed February 2024.

the Research Council and its related research and data centres. At this time, the Council were seeking to understand the role of different technologies in helping them to transform their data curation provision. They were interested in firstly understanding the context and problems faced by research and data centres, before then assessing the potential application of technology. Virtual Labs were not considered within this feasibility study but comparison of the interview data with blockchain enables one to ascertain how features of the technologies afford trust, or not. Importantly, this work was undertaken at a time of recognised transition for the Research Council prior to any technological or infrastructural developments (March 2018). Empirical qualitative research is undertaken to understand the context of data curation and trust within environmental data science; discover the requirements and expectations held by those curating this data on behalf of data producers and consumers; and to ascertain the potential development for technology in light of these findings. The results of this are reported from general exploration of the context and the first round of data analysis (section 5.1) to specific trust affordances in the second round of data analysis (section 5.2) wherein four participants with leadership roles within four data centres were interviewed, alongside two additional interviewees (section 4.2). Within the wider research, this chapter illustrates the perceptions of data curators (figure 4.2), their perceptions as trustors and trustees, and the role of institutions as data curators serving both data producers and consumers,

5.1 The Research Case Context

The six research centres within NERC archive approximately 15 Petabytes of heterogeneous environmental data across their five data centres.[3] This includes environmental data from around the world and across all disciplines, from glaciology and geology to biodiversity measures and food production in order to answer

[3]At the time of the study in 2018.

pressing local and global challenges (Natural Environment Research Council, 2022). NERC's research and data centres are seen as custodians, acting as a "*conduit*" (Participant C) to curate data for the long-term on behalf of scientists to many research communities within academia and beyond. This stewardship is undertaken in different ways within each of the six research centres and the five connected data centres, each focusing on different aspects of environmental data. Traditionally, those who produce and submit data are also the majority of data consumers due to the specificity of the data, i.e., niche sub-disciplines of environmental science. However, with an increase in open science and transdisciplinary collaborative research, the number of data consumers beyond this traditional remit is growing, including an increase in connection to other disciplines such as economics, social science, and medicine. At the time, these research centres and data centres were independent, and the changing nature of environmental science was not perceived in the same way between all interviewees, and relatedly, their data centres. As a result of these transformations within environmental science, some participants felt that their trustworthy reputation could no longer be relied upon to the same extent it had previously, exemplified by Participant D, the head of 'Brown Data Centre' and head of a team within the related research centre:

> "*If you would have asked me a few years ago I would have said it is more on the brand and the data centre itself, that reputation is what people would trust and therefore they would acknowledge that we did a good job and they wouldn't question any aspect of the trust.*"

Trust is often invisible or 'everyday' and not questioned unless an instance has occurred in which one becomes aware of trust or one lacks suitable familiarity in order to make an assessment of trustworthiness (section 4.3). Increasingly open and collaborative forms of scientific practice potentially challenge the institutional trust placed in data and research centres, particularly that assumptions of trust and reputation can no longer be taken for granted. Yet, in a data centre where the demographics of their users had changed at a much slower pace, i.e., fewer new data

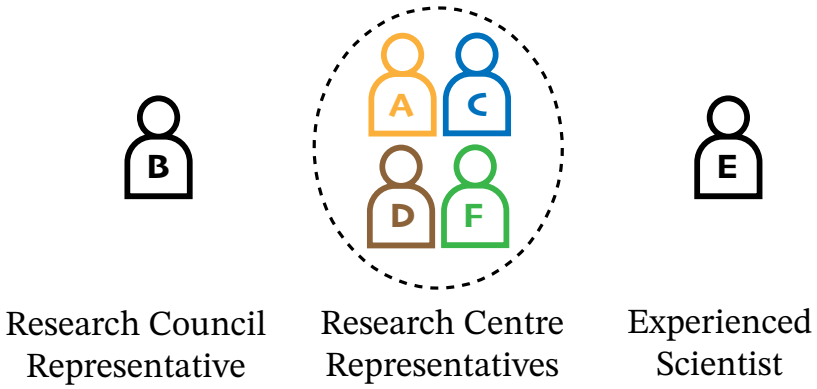


Figure 5.1: Distribution of participants and their roles. Data centre representatives are colour-coded to their respective data centre (anonymised): orange, blue, brown, and green.

consumers ('Blue Data Centre', figure 5.1) this implicit trust in the data centre — and hence it's data — was not questioned. The changes in demographics correlated to opinions on the nature of the custodian role and the potential for innovation through technology. All of the participants except Participant C, a curation manager at Blue Data Centre, had witnessed this change and felt that data centres and their provision to different users had to adapt to keep pace. Thus the nature of increasing pressures to curate data for an ever-increasing range of data consumers — and the resulting issues surrounding trust — were useful in order to explore the way in which trust was perceived by data curators in themselves and also of themselves by data producers and consumers.

In response to the shifting landscape of environmental data science, the Research Council (NERC) and its research and data centres were keen to proactively investigate the role of technology in making data available, discoverable, with contextual meta-data and appropriate licences and to understand how technology could help them to foster trust with new and existing data producers and consumers. Distributed ledger technology, specifically blockchain, was investigated to understand the role of technology in the different dimensions of trust. Blockchain

records each transaction ('block') as well as copies of previous blocks together, forming chains (Underwood, 2016). Blocks must be verified by a majority of other computers before they are added to a chain (Underwood, 2016). This eliminates the need for a controlling authority (Völter, Urbach, and Padget, 2023). Key features of blockchain are immutability, verifiability, transparency, decentralisation, and security (Brookbanks and Parry, 2022; Hoffman et al., 2018; Shin, Zhong, and Biocca, 2020). The reasoning behind the feasibility of blockchain and its relevance to scientific research and Research Councils, is for the potential use of the technology in scientific workflows (Hoffman et al., 2018), ownership rights (Shin, Zhong, and Biocca, 2020), preserving digital records (Lemieux, 2016), and providing timestamps (Elsden et al., 2018).

Distributed ledger technology is often regarded as having a revolutionary and transformatory effect on trust and collaboration (Elsden et al., 2018; Hunhevicz and Hall, 2020). With some arguing that blockchain results in a form of 'trustless trust' (De Filippi, Mannan, and Reijers, 2020) or "trust in a system without having trust in the systems actors" (Werbach, 2019 cited in Brookbanks and Parry, 2022). This 'trustless trust' is theorised to be achieved because users rely on the cryptographic nature of the technology rather than other users (Craggs and Rashid, 2019). However, many authors have argued that rather than blockchain circumventing the need for trust altogether and hence the emergence of 'trustless trust', there is rather a different form of trust (Botsman, 2017; Brookbanks and Parry, 2022; Craggs and Rashid, 2019; Ostern, 2020) that 'shifts' (Völter, Urbach, and Padget, 2023) or 'displaces' (De Filippi, Mannan, and Reijers, 2020) the traditional conceptions of interpersonal, institutional and technological trust. This has led to the study of which trust models can accurately reflect the introduction of blockchain into a context (Harz and Boman, 2019; Ostern, 2018) including the identification of familiarity, transparency, and past credible commitment (Völter, Urbach, and Padget, 2023); both technology-based and usage-related factors (Ostern, 2018); and a blend of technological, social, and institutional trust (Sas and Khairuddin, 2017).

5.1.1 Themes of Trust Relevant to Data Curators

During the original analysis of the interview data (figure 4.3), six themes were derived following the protocol outlined in section 4.3.1: standardisation; supplementary information; interactivity; provenance; traceability; and management of stakeholder interests. These themes were identified as the areas in which trust is a recognised problem for data curators and the process of data curation within environmental data science for this Research Council and its research and data centres (Thornton, Neumann, et al., 2019). Standardisation was considered a foundational challenge, that must be addressed before moving on to any other challenge. There are "*different sources of data, different ways of collecting data, different groups collecting the data*" (Participant E) that makes it difficult to "*standardise across the data in terms of its terminology*" (Participant A). This heterogeneity makes it difficult for those familiar with a sub-discipline to utilise data, but is even more challenging for those with a lack of familiarity. This also has impacts when integrating data, both in format of the data and in the aggregation of different variables.

Supplementary information in the form of meta-data, written documents, and a peer-reviewed data paper were of importance. NERC-funded projects are required under their funding to submit the data to the relevant data centre but this can be labour-intensive, with data centres having to 'chase up' data producers. The reasons for this include reluctance to submit data, lack of time, and lack of understanding of its importance for data reusers (Participants A, C, D and F). For instance Participant F noted that "*the people who get the data out do want to know these things, the people who are putting the data in can't understand why you would want to.*" Essentially, "*because they generate their own data, they don't tend to use the data centre to get data out, so we have got two different groups of people we are dealing with*" (Participant F). Three participants highlighted the importance of supplementary information in providing a narrative and context for sensors and autonomous vehicles in comparison to traditional forms of data

collection. Whilst data has traditionally been collected by scientists and there may be biases or variations within the collection process, data consumers are aware of this and can take this into account. However now with different data collection instruments, data needs to be interpreted with respect to quality, placement, and the surrounding context as this can impact the results produced and the trust placed in data. Different data centres handle this variously, e.g., Participant F noted that in 'Green Data Centre' they have, "*so many meta-data fields on our system that require information for us to actually be able to take the data in so we are quite demanding on the meta-data front definitely*" whereas Participant A representing 'Orange Data Centre' felt that, "*we haven't got a good structure for it yet. Maybe we never will, maybe it will always be unstructured.*"

Interactivity was considered important to those data centres moving towards new consumers for both discoverability, "*we have data catalogue pages but in some cases they are quite opaque [..] having to go to many different sites, and in some cases using many different log-ins*" (Participant B) and in offering a service, "*some of the really big data centres are already there. Certainly in environmental sciences we are nowhere near to be able to do that, to give people the information that would be useful to them*" (Participant A). Other research and data centres, have trialled certain forms of interaction, "*our catalogue has a mechanism attached to it so that people could write a comment next to it and we had a commenting system. Nobody used that at all, they are just not interested*" (Participant C, 'Blue Data Centre').

Provenance as a theme covers both where the data has come from, but also where the data goes to. Provenance was seen as important in providing context for data consumers. This influences their ability to use data, particularly aggregated data, which is dependent upon a level of understanding to determine appropriateness, usefulness, and to make decisions about whether to use it or to discard it. Rather than provenance as the source of data (evidential or coarse-grained provenance, section 3.4.1), the focus for three participants was on the supplementary information along the data journey providing a narrative of all assumptions, methods, and

practices made throughout (informational or fine-grained provenance). Provenance was also identified as important for data producers particularly when it comes to the use of their data, "*it is something that keeps them awake at night*" (Participant A) because of a "*risk that it is going to be inappropriately applied in the future for something it shouldn't be*" (Participant D). In their role as a curator sitting between producers and consumers, finding a way to attribute credit and provide assurance, to suit all stakeholders and maintain their curatorial duties will be challenging:

> "*In the future for the Research Council, lots of different datasets will be combined together to come up with new products, and this is where it will get really complicated to keep the provenance, and to be able to attribute credit*" (Participant F).

The final theme, reflective of the juxtaposing opinions held by participants is the management of stakeholder interests. This represents an overarching theme representative of how any technological innovation introduced will have to cater to several type of stakeholders. For data producers and curators, efficiency of the process was noted. Currently, when finalising a data upload to a data centre, "*there's a load of back and forth, back and forth, which becomes disjointed*" (Participant F). But for data consumers there is a variation in their needs. For pre-existing research and data centre users, efficiency of current processes may be needed, but others might need to provide "*support for improved customer service*" (Participant B). In terms of the role of data centres as curators, they struggle with moves towards open access with no registration, "*if we discover a problem with a dataset we have no way of contacting them other than putting out a public notice*" (Participant A). So the challenge therefore exists in managing these competing interests within system design whilst remaining cognizant not just of data producers and consumers, but also of data curators.

5.1.2 Workshop

The six themes were presented and feedback was gained on the presentation at the workshop held (section 4.3.2). Workshop participants were receptive to the themes, agreeing with the content and importantly the nuance presented. A photograph of the sticky notes collected during the workshop session and of the key notes from the verbal feedback following the presentation are included in Appendix E. No participants voiced any disagreement with the themes and their content, but added to the findings based on their own experiences.

With respect to the interview and workshop participants, the discussion around trust focused primarily on provenance. This can partly be attributed to the focus placed on distributed ledger technology. By design blockchain makes the modification of data impossible once it has been verified. This introduces reliability and authenticity. However, participants raised the issue that this does not mean that the submitted data is accurate, therefore some other technology would also be required in order to ascertain verifiability. In some cases if an error is found they would not be able to update and correct the data as they currently do, and therefore blockchain may complicate the processes to rectify this. They also mentioned a potential downfall in traceability, if data is taken from the blockchain by a user there would be a record of this but they would be unable to trace it and so would largely remain in the same position they are now. Finally, one participant in the workshop (as shown in E.1) noted that this immutability may negatively impact upon the level and detail of data submission for some users. Precisely because blockchain concretises data, this may be a factor contributing to distrust.

Overall, blockchain provided a useful example of technology to uncover and explore trust within the setting. With regards to dimensions of trust, blockchain is most applicable to trust in data (section 2.5). However, blockchain did not sufficiently cover all aspects of trust that data curators perceived as relevant to their role, with the unifying sentiment being that trust is fluid, context-specific and person-specific but conversely, blockchain is not. Whilst addressing some trust

issues, the nature of blockchain may cause others to arise. As discussed throughout the thesis, the aim of the research is not to focus on the design and specifications of a set technology but to uncover the higher-level elements of technology that may afford trust. Grounding in specific technology and it's features can provide the opportunity for insight into potential affordances between features of different systems (section 1.3). Elements of blockchain can be subsumed within Virtual Lab design, and thus the intention is not to pit one technology against the other. Virtual Labs were not considered during the initial conduct of the research and could thus not be applied to what participants in either the interviews or workshops discussed. However, during the second analysis of the data (section 5.2) and within chapter 6, Virtual Labs as a potential technology within environmental data science are considered. When arriving at specific instances of trust affordances related to data curators' perceptions of trust in the next section, the opportunities that may be afforded are not specific to either Virtual Lab or distributed ledger technology, but could be applicable to any context in which technology has the features to enable them.

5.2 Developing Trust Affordances

The perceived problems in fostering trust for data curators signal the need for a range of trust affordances. Three categories were identified from the social computing literature and outlined in section 3.4 in order to draw out general classifications of trust affordances. As per the research design (figure 4.3), the qualitative data collected in the interviews for this case were re-analysed in order to create specific instances of trust affordances, codes were created and grouped in order to build up individual trust affordances and compared with the general categories iteratively, neither inductive nor deductive. These data curation trust affordances for this case are presented in table 5.1.

Category	Trust Affordance
Provenance	Immutability
	Verifiability
	Traceability
Transparency	Scrutability
Social Translucency	Visibility

Table 5.1: Identified trust affordances and their related category identified within the Research Council case.

5.2.1 Provenance-Related Trust Affordances

Participant D noted that some users of their data centre "*will sub-sample many different datasets in the course of coming up with a new product via some algorithm or hypothesis, they have seriously got to be able to understand the provenance of the information that they are using both actively or discarding from what they are generating if they are going to have any understanding of the risk and uncertainty.*" As discussed previously, provenance can take multiple forms.

Immutability refers to the persistence of data, that it will remain in the same location unchanged once it has been submitted (Craggs and Rashid, 2019; Ostern, 2018). This affordance provides an opportunity for one to trust that whatever was written originally, whether spatio-temporally distant or not, is the same now as it was then. A recurring instance within the interviews was that some data, particularly older data, does not have a record detailing where it came from and with associated meta-data. Several of the data centres (Blue, Brown, and Green) noted that they have started producing a Digital Object Identifier (DOI) for datasets, akin to those produced when publishing academic papers "*in order to give people a little bit of trust that there is actually- that the datasets are going to stick around and not be changed so much*" (Participant C). Yet, whilst this method is used and can be a source of encouragement for data producers to submit data because of newer requirements set out by publishers to have a DOI for data, some participants did not

have positive remarks to make. For instance, Participant F described the process as a "*crude version of blockchain*" because they had experiences of both real-time data (with value for some applications, e.g. weather data but known to have lower accuracy) and frequent changes to submissions wherein "*somebody changes the data, they come along and say 'oops, we put the wrong calibration on' then you have to deprecate the existing copy and now starting using the new copy, which will refer to the old, deprecated copy and say 'there used to be this copy and it's been deprecated'.*" Thus whilst this method may be appropriate for some data, this is not a full-coverage solution in all cases.

Other participants discussed the idea of immutability as having negative impacts. Participant E noted that "*there is perhaps a feeling amongst some that there would be comfort in knowing that something is static, i.e., that the record wasn't going to change. The trend now is to have a living document. The trend is going in the opposite direction.*" Similarly, Participant A spoke of the same trend in regards to data where it is "*not seen as a static thing, its seen as more of a living resource that you can add value to and explore different nuances.*" This speaks to the purpose of technology and perceptions of it playing a role in the perception of immutability as a multi-sided affordance, for some having a permanent and persistent record is a positive based on their beliefs of what purpose a technology should have, whereas for others, the fact that it can be changed or added to is a positive, albeit introducing more work for data curators. Beyond the immutability of data, there is also a focus on the related supplementary information and the immutability of this, i.e., that it will remain connected to the data throughout it's journey.

Verifiability is a recognised affordance of information (Eriksson and Eriksson, 2019; Weltevrede and Borra, 2016) and contributes directly to trust in data (sections 2.5 and 3.4.1). With no "*contextual meta-data here, we may utilise some of this information in the future without understanding all of the issues that go around with that*" (Participant D). This version of verifiability speaks to provenance-as-context. Fine-grained, 'why' informational provenance provides insight through meta-data

and documentation. However, "*that kind of provenance information is quite difficult often to convey to the audience because they have different needs in terms of how they want to see that and what types of information they want*" and can impact on "*whether they see that as a valid dataset or not*" (Participant A). This relates to the ability to verify data as trustworthy by data consumers in their own terms, i.e., whether it is valid and accurate and thus reliable.

Verifiability can also relate to provenance-as-source. This is most relevant to issues of legality and proving ownership. Coarse-grained, 'where' evidential provenance affords insight into the identification of the data producers or instruments used. For instance in the case of hydraulic fracturing ('fracking') one participant noted the difficulty in providing data for a contentious political issue that is scrutinised by various stakeholders. They stated that providing verifiable evidence of sensors and their measurements as a 'true record' will be essential for their role as a data curator.

Traceability is recognised as an affordance (Mahr and Huh, 2022; Ostern, 2018; Tabuenca et al., 2021) providing technology users the opportunity to track elements, i.e., data, that they are concerned with and thus enabling agency and control. As with the other named provenance-related trust affordances, traceability is not one-sided. For data consumers, Participant E described this as challenging, regarding one project they worked on:

> "*We had to manually unpick the data, and a lot of the information was lost [..] there is a number here but 'what was it derived from? What was behind it? What models were run? What assumptions were made? Who ran these models? Are they trustworthy?' All that was lost [..] and then we started to find numbers we didn't believe, and we had no basis to check why, this number looks wrong but you have nothing to go on [..] apart from going back to the source if you are lucky enough to have that.*"

As data curators, data centres have tried to aid this traceability. Participant D noted that they "*were hoping*" that the DOI system would work to "*trace citation*

of whole datasets, albeit then to a project that tells us it discarded most of that data and didn't use it anyway..." The past tense and offhand remark regarding the use of DOIs may signal that the process did not work in the way they had theorised.

Traceability can introduce negative impacts as well. It is linked to identifiability which can lead to distrust (Ostern, 2018). As discussed previously, identity can afford trust for some users in achieving credit or determining reputation, but for commercial users there may be confidential or sensitive data that they don't wish to be connected to. Thus striking a balance between what is traceable, to whom, and who has access this to this information is essential in determining how this trust affordance impacts the various stakeholders a data centre caters to.

5.2.2 Transparency-Related Trust Affordances

As an affordance, **scrutability** relates to the opportunity to scrutinise or examine in order to make comprehensible (Kay and Kummerfeld, 2010; Noriega et al., 2015). This can relate to data and models or a system and speaks to cognitive trust (Passi and Jackson, 2018). This affordance relates to the balance of openness vs. privacy. Two participants (A and B) spoke about the low levels of scrutability currently existing within data centres characterising it as opaque, often incoherent, and high-effort. It was felt that these characteristics precluded users from benefiting from the data a data centre could offer, exacerbated by the variety of environmental science data available. From a data curator point of view, they felt that in being scrutable, they could cater better to data consumers. By enabling users to investigate or uncover data for themselves, scrutability would be improved. This is sensible as accessing data or outputs can be frustrating, time consuming, and ineffective (Dean et al., 2017; Lee and Stvilia, 2017). Yet, curators also recognised that this would not work for every data consumer. Participant A spoke of a dynamic and user-specific system because "*rummaging through lots and lots of different datasets and then downloading the one you want is not going to be practical. What you are actually wanting to do is engage in a dialogue with that data centre through some kind of*

service interface." They perceived this system as having the ability for users to specify their needs, "*I'm interested in soil moisture data over the last forty years, and I'm particularly interested in south west England, what have you got?*" and then the system to suggest either a specific dataset or a combination of data sources to fulfill the users needs, to which the user could agree or disagree, and then "*it would then do whatever it does behind the scenes, to give you that bit of data*" (Participant A). Clearly, this enables srutability, but is a different form to "*a one stop shop or a single point of entry*" (Participant B), which could also aid scrutability by different means.

5.2.3 Social Translucency-Related Trust Affordances

Social transluency speaks to the ability to undertake communicative and collaborative practices, afforded in multiple ways (section 3.4.3). **Visibility** sheds light on the information that we are usually 'socially blind' to in the digital realm (Erickson and Kellogg, 2000). With respect to the interview data, three participants spoke about the connection between identity and trust. It speaks directly to non-technological affordances as described by Norman and Gaver, e.g., the opportunity to determine through sight whether a door is a push or a pull (section 3.2). Two participants, A and E, mentioned that reputation of scientists and data producers was useful in determining whether or not to trust data, for instance Participant E noted that, "*even if it's 'just scientists' there's still scientists you trust and scientists you may not trust.*" Both participants inferred that who the person is would have an effect on their perceptions of the trustworthiness of data and that the absence of this information would make the determination of trustworthiness much more difficult. Participant D built upon this, by connecting the known identity of a contributor as an opportunity for collaboration and accountability (as found within the literature, section 3.4.3). As discussed above, this visibility may conflict with other trust affordances, dependent on the specific instantiation it takes. For some, visibility leads to openness and honesty (Clark, 2014; Turilli, Vaccaro, and Taddeo, 2010),

refocused as a form of embeddedness, "contextual properties provide incentives for the trustee to behave in a trustworthy manner" (Riegelsberger and Sasse, 2010, p.5). The affordance of visibility can persuade trustees to act in a trustworthy manner precisely because either they or their actions are perceptible. This can deter potentially untrustworthy acts, but is not a complete solution to affording trust (O'Neill, 2018). Viewed in these terms, the same opportunity to shed light on whom one is interacting with or using data from is perceived either positively or negatively and may require some 'fine-tuning' or advanced functionality to turn on or off this feature or features.

5.3 Chapter Takeaways

The research undertaken within this chapter occurred during a transitional period for the Research Council, its research centres and data centres. Recognising the fact that all within the NERC umbrella must adapt and shift their data curation provision in the face of changes in scientific practices, the Research Council was seeking to understand the feasibility of different technologies in order to improve their offering as curators to those within and beyond environmental science. In the face of changing demands, they felt that they could no longer rely as heavily on the elements that traditionally made them feel perceived as trustworthy, and that they must also transform. As discussed at the outset of the chapter, the empirical research presented here focused on distributed ledger technology. Virtual Labs at this moment in time were not considered. It should be noted however that since this time, Data Labs — a physical instantiation of the Virtual Lab concept — have been adopted and continue to grow (Hollaway et al., 2020). And, as noted at the beginning of the chapter, the creation of the Environmental Data Service — a combined curational facility — has also been set up.

This chapter details the exploration of trust amongst data curators, those responsible for preserving and promoting access to environmental data and models.

Within the case, trust was considered an important element within environmental data science to varying extents. The majority of interview and workshop participants recognised the role of trust and were able to communicate the reasons behind this and to identify specific elements of trust that they considered to important. Notably, one participant (Participant C, 'Blue Data Centre') did not consider trust to be a concern for themselves or their data centre users but upon further investigation, they were able to expand upon their opening position, "*I think if you asked somebody 'do you trust this data?' people would- our users would, interpret that as 'is this the right quality?' not 'where did it come from?'. That's what they would go straight to, quality.*" However, contrary to most participants, Participant C did not connect the source of data to trust. Within the design of a Virtual Lab, supporting institutions (such as NERC and its affiliated research and data centres) will impact upon the levels of trust perceived by stakeholders (chapter 2). Thus an important contribution is not only the perceptions of data curators themselves, but also of the wider institution responsible for data curation, hence the inclusion of Participant B who could speak to the wider aims of the Research Council as a whole. Data curators were open to the need for fostering trust and considered this an essential part of their role.

Affordances are common within the computing literature, and many of the identified trust affordances exist within their own right or are identified as opportunities but not labelled as affordances by their respective authors. For instance, identified affordances of blockchain include privacy, security, transparency, and accountability (Shin, Zhong, and Biocca, 2020) and accessibility, immutability, responsibility and traceability (Craggs and Rashid, 2019; Ostern, 2018). To date these terms have not been analysed in relation to affording trust and thus this research presents a new understanding for the design of technology to promote well-placed trust (Thornton, Knowles, and Blair, 2021). With respect to trust affordances, of the total identified from the empirical data, provenance was the highest out of the three primary categories (table 5.1). Owing to the fact that this

case is based upon a feasibility study investigating the role of blockchain in fostering trust, most participants did think primarily of provenance, focusing on where data comes from, and it is possible that the nature of the interviews and workshop may have led some participants to privilege this aspect of trust despite neutrality from the researcher. The trust affordances presented within this chapter pertain to the perceptions of the data curators consulted. Given the treatment of trust thus far, i.e., a subjective, idiosyncratic, and contextual phenomenon, these affordances will be contrasted with the affordances of data producers and consumers derived in chapter 6 as a basis for discussion (chapter 7) with a particular focus on how differently trust affordances appear based upon context. The same affordances may or may not be identifiable or present opportunities for different groups of stakeholders, and thus may have implications when taking a trustworthy-by-design approach.

Chapter 6

Exploring Data Producer and Data Consumer Perceptions: Trust within a Centre of Excellence

This chapter considers the perspectives of experienced environmental data scientists as data producers and consumers, grouped by their roles as theme leaders within a Centre of Excellence. The Centre was launched in 2019, formed as a collaboration between a University and a NERC research centre. Those interviewed represent several sub-disciplines of environmental data science, represented in figure 6.1.[1] Following the research design (figure 4.3), results are presented for both the first (section 6.1) and second (section 6.2) rounds of data analysis, in which the data is analysed for general themes related to trust and then later, to identify trust affordances to aid in system design. In comparison to the case in chapter 5, this case focuses on the perspectives of those both producing and consuming data, and

[1]Although it should be noted that this separation was created by the Centre in order to structure their leadership and in reality the majority of participants expressed experience with both environmental specialisms and methods rather than one or the other.

thus interacting with all types of stakeholders (producers, curators, and consumers) as depicted in figure 4.2.

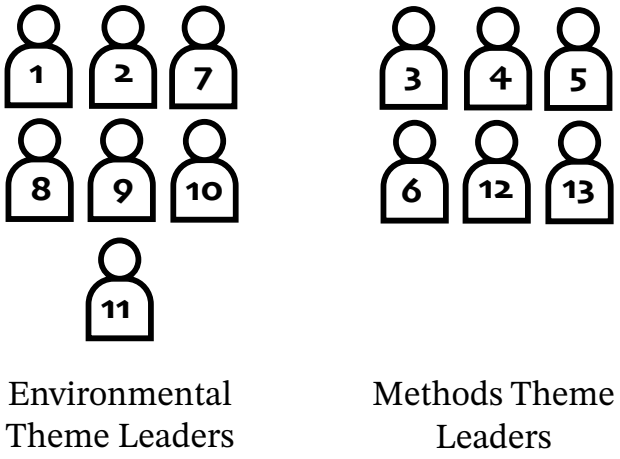


Figure 6.1: Distribution of participants and their specialism as theme leaders within the Centre of Excellence. Environmental themes include: air, biodiversity, land and soil, and water and ice. Methods themes include: data acquisition and infrastructure, data science methods, decision-making under uncertainty, and complex systems. Participants are labelled with a number for anonymity.

6.1 The Research Case Context

All of the participants are either environmental scientists by background and currently working in their specific environment sub-discipline or are academics who work within other disciplines, e.g., computer science, mathematics, and statistics who work on projects related to environmental data science. Participants were receptive and positive to the study of trust within environmental data science. Trust and trustworthiness as concepts have not been paid attention to "*historically*" (Participant 8), "*but will perhaps increasingly come to the forefront because more and more decisions are going to be made based on environmental data*" (Participant 4). Two general trust themes were identified during the first-pass of data analysis,

wherein codes of the data were compared and grouped. Two themes comprising *social* and *technical* aspects of trust were identified with two sub-themes within each theme. These two themes were developed based upon participants' descriptions of how they determined to place trust, through social and/or technical means. Metaphorically, one can imagine a yin and yang comprised of social and technical factors (a preliminary drawing to illustrate these themes is included in Appendix F), symbolising the advantages within one theme helping to complement shortcomings of the other and vice versa. For instance, many participants focused on the technical role of documentation in providing supplementary information, but also recognised the aspect of the social, including the ability to have conversations when documentation did not provide them the insights they required. The strength of these themes varied along the spectrum, from Participant 8 who stated that "*it's not so much trust in things, as trust in people*" to Participants 4 and 12 who stated, respectively, that trust in data is "*always an issue*" and "*you've definitely got to trust, trust your data source, trust your methods.*"

6.1.1 Social Aspects of Trust

Social aspects refer to the categorisation of trustworthiness cues related to people, whether this be interpersonal or impersonal, i.e., through an insitutional affiliation. Participants were not given the theme and asked to speak on it, rather two sub-themes of *ability* and *integrity* were created when grouping codes based on participant discussions across the three dimensions of trust (trust in data, trust in people, and trust in systems).

Ability was identified as a source for placing trust. Focusing on the perception of skills, competence and knowledge to complete a task (Mayer, Davis, and Schoorman, 1995) such as data collection "*we have a fairly reliable evidence base that the people who are collecting the data do a good job*" (Participant 8). The ability of data producers was seen to impact on the quality and the appropriateness of data through their ability to plan and control data collection methodologically. If this is not done

correctly or suitably, the resulting data may not be worthy of trust. Speaking about their own ability in comparison to others, Participant 8 noted that "*I don't have any field experience. So I wouldn't go out and collect data. I go with somebody who has field experience, so they will be cognizant of the limitations and the challenges that are in place in terms of collecting high fidelity data.*"

This notion of ability also reflected certain types of data, such as that collected through citizen science projects. Participant 1 contrasted the different abilities of those they collect data from. In the first instance, they described their relationship with a field scientist to identify species:

> "*There are issues in that one around who's recorded it because some of it is so subjective, especially with things like which species [..] it's not always clear. And, and so for me it works because I know the guy who inputs a lot of the data and he's one of my collaborators and so we can discuss that, but I'd be really nervous about using it if anybody else has put the data in.*"

In comparison a proportion of the data Participant 1 uses comes from a database where:

> "*It's probably thousands of people. And because you can, anyone can upload anything basically, so it goes from people that are kind of amateur naturalists who will have seen birds in some location, and they upload that data to people doing kind of massive research projects who are uploading it. So there's a huge range of people that are putting it in there.*"

They stated that transparency, honesty, and reproducibility were essential in using data from these large-scale databases in order to determine quality and appropriateness of the data for their own work. This was also reflected by Participant 6 who stated that transparency was something that they included in their online research environment in a previous project, when they found that users might ask, "*what is behind this? Who's coming up with the data? How trustworthy is this data?*

What are the models being used? Is this model sound? Is this something that we can trust or is it you know something that you know an undergraduate knocked up in a couple of evenings or something, you know?"

Integrity was identified as the second social sub-theme during data analysis. This is the perception that a trustee adheres to principles, has consistent past actions, or has a strong level of reputation or recommendation from others (Mayer, Davis, and Schoorman, 1995). Trust was identified with regards to collaboration. For Participant 3 this involved trust-building as part of the collaborative process, whereas for Participant 8, they collaborated primarily with those that they already trusted. These collaborative relationships helped them to share and receive data (Participant 1) and to perceive research as trustworthy, "*I think that is where you get the element of trust, it's the fact that it's not just one person blowing their own trumpet, it's a group of people that have worked together*" (Participant 12).

Reputation was explicitly identified as a signal of trustworthiness by four participants as a marker to trust in data. This was exemplified by Participant 8 who stated, "*reputation is a big one. So scientific reputation, someone who's well respected in their respective fields. So, you don't necessarily have to know the details of what they do. But you have to know they're good at what they do.*" There was appreciation that as experienced scientists, their experiential knowledge of others' reputations would not extrapolate to those earlier in their career, and that cues of trustworthiness are therefore dynamic over time. Participant 8 neatly summarised the issue:

> "*The nature of the work we do in the environmental data science space is very collaborative, because I don't have the statistical or computational abilities and skills that are needed to do some of this work. So I have to do it in collaboration with experts [...] so there's definitely an element of trust there, in that I trust that [name] knows what he's talking about from a [discipline] perspective. [...] it's easy to do that with [name] because I know them quite well, and I've known them for a while and I have a lot*

> *of respect for them as scientists. But I can see how people starting up new collaborations, it could be quite difficult.*"

Integrity requires some knowledge by the trustor of the trustee in order to determine whether there is consistency with past actions and their values. Facilitating trust using integrity cues will likely require different solutions, i.e., recommendations, and the promotion of relevant cues of trustworthiness to complement this. However, from a trustee perspective, Participant 11 raised the issue with the integrity of scientific research and negative outcomes:

> "*I think that the communication to the general public is much, much more difficult because there is always then the risk of sounding too certain and then losing that trust in scientific evidence versus not sounding as though we have a clue what's happening [laughs]. So it's always a balance where you have got to be open and honest about the uncertainties but you also have to be careful about the headlines that get picked out, without all the other stuff that is surrounding it.*"

6.1.2 Technical Aspects of Trust

Technical aspects of trust were identified during data analysis, consisting of two sub-themes: *detailed information* and *data quality* (Appendix F).

Detailed information, including meta-data and additional supplementary information, was identified as important for understanding the data and the surrounding context by all participants. This was regarded as essential for using data, but a recurring problem in practice. Despite the paramount importance placed on detailed information as a data consumer, several participants were open about the difficulties they had faced as data producers. This included: lack of space to reflect on and publish the details of the data collection and analysis (Participant 13); the anonymity of the end user, which limits how to tailor documentation (Participant 4); a lack of time and motivation to fully document everything (Participant 13); lack

of standardisation across sub-disciplines regarding what counts as tacit knowledge and therefore doesn't get communicated (Participants 6 and 8); and a recognition that their own notes would be unintelligble to those they have shared data with (Participant 13).

Data quality impacts on the ability to use data and models. Participant 9 stated that when using data collected by others "*without seeing all of the, the QA [quality assurance] work and validation that's been done in some of the other models, I would have had very low confidence and trust in their outputs.*" Data quality was primarily linked to data source by participants, e.g., it was most frequently mentioned when considering data from repositories (Participants 1 and 5) and from citizen science projects (Participants 2 and 5). Quality has to be balanced with availability. Two participants (1 and 8) noted that because there is often not enough data for their research, they have to make assumptions and decisions about whether to use it or not. This decision will be impacted upon by other themes discussed, i.e., if there is no detailed information it will be harder to ascertain quality.

Both of these technical sub-themes cover different aspects of participants' experience as a data producer and a data consumer. As a producer, it is challenging to provide information to a wide and unknown audience, yet as a consumer, this is central to forming trust in data and models which one has not collected or created. This is related to the previous case, wherein data curators expressed a tension between encouraging data producers to provide sufficient detail that data consumers want when they cannot see the benefit as non-consumers (section 5.1). Those interviewed within the Centre of Excellence case rarely mentioned the specific role of data curators, despite a large number of them being employed by a research centre with a related data centre. For instance Participant 5 spoke about citizen science data and the caution that they exert when utilising this data type, "*it could still be the case obviously that the data from that is then in some sort of reputable source like the data centre*" hinting that despite the reputation of a data centre, the data hosted by them is not automatically trustworthy. This tension between

stakeholders and the ways in which they determine the placement of trust exists across cases reiterating the cultural shift needed within environmental data science alongside technological development. A holistic approach is needed to communicate between groups prior to instantiation. Trust, being so idiosyncratic, may make it difficult to view from others' perspectives thus feeding into the difficulties illustrated within this research. Features that facilitate trust for one stakeholder group may impact the trust of another group.

6.1.3 Survey

Themes derived from the interview data were presented in a seminar and a link circulated to attendees and the wider Centre mailing list. Half of the survey participants (n = 5) described themselves as a post-doctoral researcher, with the remainder being a doctoral research student (n = 1), an academic lecturer or above (n = 1), and a technical developer (n = 1). When presented with a Likert scale regarding the statement "*Trust and trustworthiness is important for environmental data science as a whole*" 8 participants selected strongly agree and 2 selected agree.[2] A follow up question asked whether "*trust and trustworthiness is important for me in my research and/or role*" 6 participants strongly agreed and 4 agreed. Bar charts presenting the results of all quantitative questions can be found in Appendix G. [3] It should also be noted that one of the participants dropped out after answering the first five questions and for some of the questions participants were permitted to skip answering if these were not relevant to their research or role.

On the whole, participants responded that they had considered trust previously, which indicates that there may have been some instances in which trust had been brought into question and examined. This sets the scene for the remainder of the

[2]Where 1 = strongly agree, 2 = agree, 3 = neither agree nor disagree, 4 = disagree, and 5 = strongly disagree.

[3]These charts were created by exporting the data from Qualtrics as a CSV file and uploading the data to Microsoft Excel.

survey, which sought to understand the ways trust is important.[4] Regarding the themes derived from the interview data, the majority of survey participants felt that "*trust to me is associated with a mixture of social and technical*" factors (n = 6), with 3 respondents associating trust primarily with data, models, and outputs and thus technical, and no participants feeling that trust is primarily associated with people and relationships. Free text answers expanded upon these views, showing an appreciation of robust, rigorous, and thorough reporting in order to ascertain the trustworthiness of data and to understand uncertainties, reliability, and accuracy. Alongside documentation to support their assessment of trustworthiness, respondents also mentioned peer review, reputation, use of models and data similar to citation metrics, and communication with others who have expertise.

Regarding specific mechanisms of trust development within environmental data science, there was support for knowing provenance (6 = strongly agree, 3 = agree), awareness of identity (2 = strongly agree, 2 = agree), and the use of documentation in the form of meta-data or supplementary information (6 = strongly agree, 2 = agree). Although notably, six out of eight participants neither agreed nor disagreed that they have had smooth, positive experiences of using documentation. When thinking about translating the work into Virtual Labs, there was mild support for modified access through credentials (2 = agree, 6 = neither agree no disagree, 1 = disagree), strong support for transparency (6 = strongly agree, 3 = agree), and support for communication features (2 = strongly agree, 7 = agree). There was support for the ability of Virtual Labs to improve collaboration (2 = strongly agree, 7 = agree). The results of the survey confirm the findings derived from the interviews. Survey respondents agreed with the points raised by interview participants, and importantly, reflected the nuance and spread of reflections on both social and technical aspects of trust.

[4] As mentioned previously, trust is often not questioned unless an instance has occurred in which one becomes aware of trust (sections 4.3 and 5.1). Hence, that participants have been cognizant of trust signals that they have previously considered it on at least one occasion.

6.2 Developing Trust Affordances

The empirical data is analysed in a second-round in order to develop trust affordances across three themes (section 3.4), following the same protocol as the previous case (section 5.2). Seven trust affordances are developed from the Centre of Excellence case by interpreting the statements made during the interviews, as shown in table 6.1. These trust affordances will be compared with the trust affordances from the first case in chapter 7.

Category	**Trust Affordance**
Provenance	Traceability
	Verifiability
Transparency	Scrutability
	Abstractability
	Reciprocality
Social Translucency	Perceptibility
	Comprehensibility

Table 6.1: Identified trust affordances and their related category identified within the Centre of Excellence case.

6.2.1 Transparency-Related Trust Affordances

Based upon participants statements, **scrutability** is identified as a valuable trust affordance. Capacity to understand through examination was deemed important. For Participant 5, "*the detail is hidden a lot of the time, it's very hard to get at, that sort of level of detail. [..] That it's very hard to sort of- with third party data and particularly historical data to sort of understand decisions that were made at the time how the data was derived entirely. Whilst you may have a good idea, and a general understanding, the detail is often hidden, missing, or not exposed*". Participant 13 also expressed a similar opinion, "*anything that you do with the science, you have to*

understand the process from the data through to the results". Later in the interview, they expanded on this point:

> "*And quite often, actually, the more you know the topic, the more critical you are, rather than the other way around. [..] when I can't find a way of interpreting it all, if the maths is too complicated for my brain to get round, I struggle to trust it. That said, if somebody brighter tells me that it makes sense and it's all fine I can be persuaded to just go, 'Okay, fine. I can't.. but you say it's fine. Okay' [laughs]. It depends how much I need it and how much I need to personally understand, how much it feeds into the things that are coming out of my stuff*" (Participant 13).

When this scrutability is lacking, other avenues to develop trust or determine trustworthiness are explored, though as discussed previously, this may not always exist as a possibility. For scientists as data consumers, the opportunity to scrutinise and delve into the surrounding context is important in order to utilise the data and to develop trust in one's own results. It is important to note that, at least for the scientists interviewed within this case, honesty and integrity are derived from scrutability. When data or a model is inscrutable, one cannot trust. Participant 12 noted that, "*I think it makes people trust, oddly, if you quote uncertainty on your data, I think it makes people trust your data more, because they don't think you're trying to hide anything*" with similar sentiments echoed by Participants 2 and 4.

Abstractability concerns the provision of different levels of information about the same dataset or model for different levels of user knowledge, with similar documents combining different levels of abstraction or detail. It is not commonly used as a term, but has been used within communication literature pertaining to the transfer of knowledge (Li, 2007; Vakarelov, 2012). For expert users, a high-level of detail will be required, evidenced by Participant 12, "*I find that's where I gained the most trust in a particular method if I can actually open the bonnet and start looking inside the engine and see what's actually driving it*". Participant 10 also vocalised the same thoughts, stating that "*as a scientist, you do want quite detailed meta-*

data" whilst recognising that other stakeholders may not need this information but "*knowing that the information is available will already probably garner a bit more trust than it being sort of very closed off.*" For non-expert users (or experts without the specific sub-disciplinary background), a significant amount of work is needed to design the types and levels of information presented. Participant 6 spoke about their role as a computer scientist designing a virtual research environment focused on flooding. During the course of their project they had to design the system to produce information for a range of stakeholders, including environmental scientists, Local Authority decision makers, local communities and, insurance companies. To facilitate abstractability they built web-based tools based upon typical questions that general users might want to answer with the opportunity for more detail-orientated users to adapt these"*pre-set formats which are accessible for the general public*" by changing model characteristics such as how the land is farmed and the effects of different land management activities. However, Participant 13 raised the issue of the time and motivation as a data producer to compile different abstractions of information. In their own research project they noted that, "*there's 20, 30, 40 pages of comments... And those comments, I mean, you could read them. Somebody else could read them and they would be fine. And they show the messiness of the decision. And I think that's brilliant as a way of recording something because I think that is the reality of the mess, but who's coming in from outside that's been given a dataset wants to sit and read 40 pages of us arguing backwards and forwards?*" Abstractability speaks to the demonstration of honesty and integrity whilst catering to the needs of different stakeholders. However, knowing the audience is challenging. Considering abstractability as a multi-sided affordance, a potential tension may arise. For data consumers a clear benefit would be tailored information given some input from them regarding their knowledge or role, with this provision being ordered and well-thought out, but for data producers this could act as an impediment to uploading data. In the case of Participant 13, they noted that they lacked sufficient time to add comments to their work, let alone produce different versions.

Reciprocality affords trust through the transmission of information. Recipient design highlights the interactive nature of communicative practices (Dombi, Sydorenko, and Timpe-Laughlin, 2022; Newman-Norlund et al., 2009). Rather than information travelling through mediators, technology — when designed with the features to enable it — can instead facilitate direct forms of knowledge-sharing (Carroll, 1996). Asking questions, and having answers can promote trust (Calder et al., 2018; O'Hara, 2019). Participant 13 stated that, "*What you really need instead of meta-data is access to the modeller. Um, so that modeller needs to be immortal, constantly working at the same organization for an incredibly lengthy period of time, have the free time to talk to anybody who wants to you about the stuff [laughs]. But the best thing you can do is to build relationships with people and then hope they never go anywhere. Which isn't long term or practical.*" Similarly, Participant 3 compared the reciprocality of current documentation to real-world contact, "*I would love to put a video camera in front of various people before they retired [..] I think we're still back in the 20th century idea of, you produce this paper and it has a dataset which has some dry, technical page and that's how people gain trust in things. That's not how people do it at all.*" When asked about features for communication, three survey respondents proposed different solutions for Virtual Labs (Appendix G):

- "Message boards and bulletin boards could be useful. No direct messages as virtual labs are all about transparency?"
- "Direct messages, specific message boards, opportunities for comments e.g. with code commits, or model/code readme"
- "Being able to contact the original poster is crucial"

It is clear that tools and features should be designed to improve transparency via reciprocality, in the hope of fostering trust.[5] However, as with abstractability, Participants 12 and 13 made clear that any additional work should not rest with the data producer. A culture shift wherein a community generates answers to questions

[5]Features and functions, in addition to affordances, will be discussed in chapter 7.

may also be necessary, yet it is difficult to fully represent and encourage the social component within a socio-technical system.

6.2.2 Provenance-Related Trust Affordances

The most identifiable provenance-related trust affordance within this case is **traceability**. The ability to follow the data science pipeline from data acquisition to analysis is needed for some environmental data scientists in order to trust (Participants 3 and 12). This is particularly pertinent given the often complex flow of data that has been combined with many sources (Participants 2, 9, and 10). Traceability helps to inform the assumptions made about the data and the processes that the data has undertaken on the journey from data producer to data consumer. This is particularly important for data that could contribute to laws, for instance Participant 4 spoke about data surrounding air quality, where sensor data might be used to inform an ultra-low emission zone, this may need to be traced back to specific sensors and measurements. For data consumers, a traceability affordance could also enable discoverability. As with the previous case (section 5.1), this information may provide cues of trustworthiness. Participant 12 spoke to this, envisaging a scenario in which a Virtual Lab user could view a dataset or model, documentation, interactive notebooks, and derive from this the applicability and viability for their research.

This also connects to the trust affordance of **verifiability**. "*It's becoming more and more important now because of the whole climate change emergency, if you're going to make really important decisions [...] you have to be pretty sure that the data you're offering and the analysis you're offering really supports that*" (Participant 4). Any conclusions need to be verifiable, i.e., demonstrably justified at every level and stage in order to garner perceptions of trustworthiness (when it is trustworthy). In the same way as the previous empirical chapter, this verifiabilty is deemed to be of the utmost importance when it comes to political or legal emotive issues that touch broad swathes of society, for instance, three participants noted their research feeding into the work of the Environment Agency, IPBES (Intergovernmental

Science-Policy Platform on Biodiversity and Ecosystem Services), and the IPCC (Intergovernmental Panel on Climate Change). Yet, verifiability was also considered at a fine-grained level. Participants 3, 6, and 12 spoke about reproducibility as the ability to see what someone has done, understand why they were doing it and how they did it, test and replicate the original finding, and then apply it in their area of research. They saw Virtual Labs as providing this verifiability, by replicating the conditions under which the experiment or analysis was run including the same software version, data, methods, and protocols to achieve coherent, consistent, and verifiable results.

6.2.3 Social Translucency-Related Trust Affordances

In order to collaborate, there are different trust affordances identified within the data. Bearing similarity to the previous empirical case within a Research Council, many participants noted that there are social barriers impeding their ability to trust and collaborate with others. This was particularly acute when it came to language and terminology. Thus a trust affordance to promote shared understanding and social translucency is **comprehensibility**. Inspired by the Shannon–Weaver model of communication (Shannon, 1948), noise effects communication flow between a sender and a receiver. Within the empirical interviews participants noted that "*you've almost got people kind of talking different languages*" (Participant 10) with terms such as 'model' connoting either a statistical model, process model, or conceptual model (Participants 8 and 10) or 'data' being an observed measurement or the output of a model (Participant 7). A lack of a common language precludes developing a common understanding (Participants 9, 10, and 13). And, this is particularly true of interdisciplinary research, wherein different disciplines define terms variously (Bauer, 1990; Bridle et al., 2013). Akin to the issue of standardisation (section 5.1), comprehensibility is predominantly a social issue that must be considered prior to technological development in order to design methods of opportunity for comprehension by multiple stakeholders into the system from the

outset.

Drawing upon the principle of social translucency developed by Erickson and Kellogg (2000), the provision of visible social information within a system leads to (in trust affordances terminology) **perceptibility**, and awareness of others' actions. As discussed within the trust affordance of visibility in chapter 5, identity provides some of this ability to perceive with whom one is interacting and conversing or utilising their work. Perceptibility builds upon this, for instance Participant 12 envisioned, "*you're in the same space, so you can see what everyone's done. Or you could just say, 'Well, I'm working on cell one, you work on cell two and vice versa'. [..] it's that co-location that is key in coherency, as well. So everyone's working from same data, same environment, same method.*" This perception of where others are, either in real-life or virtually, helps us to adapt our physical and verbal actions (Erickson and Kellogg, 2000). This affordance creates the possibility of collaboration and also opens up opportunities for accountability and openness. The translation of this into a Virtual Lab, or any technological system remains challenging, as discussed previously with the tension between visibility and privacy. Participant 6 felt that having open forums "*would alleviate a lot of trust concerns*" whereas Participant 3 considered the issue differently:

> "*I guess with a collaborative space the idea is that eventually outputs of that collaborative space will be publicly available, so you do have to worry a bit about the public face of this, but initially, the trust is partially enabled by the fact that it's a secure space and not anybody [..] can crash into this space and do stuff, you have to have credentials if you have to log in and just because you've logged in doesn't mean you can see everything. You'll only be able to see what's assigned to you, so it's a trusted space, but it is a collaborative space where people are developing things and the team as a whole should have sight of how things are being developed and have the opportunity to say whether they think that's a good thing or a bad thing.*"

Being able to perceive who is around and what they are able to view thus has potential impacts upon trust, including the opportunity to be open and honest – and potentially vulnerable – about their work. If they are identifiable they may be held to account (Erickson and Kellogg, 2000; Nissenbaum, 2001) by whomever has access, which may have an impact upon their levels of comfort as well as their trust in both others and the system itself. Having an open forum as Participant 6 suggested may be suited to a final public version of scientific work, but when research is still in progress, Participant 3 noted that exclusivity of access may help others, once again highlighting the pervasive tension between privacy and visibility.

6.3 Chapter Takeaways

The study of a Centre for Excellence represents a unique snapshot upon a collective group of experienced environmental data scientists comprising of those trained in environmental science, or a sub-discipline, e.g., ecology, alongside those who are not academically trained or hired as an environmental scientist but are experts in their own field, e.g., mathematicians or computer scientists, working amongst each other in transdisciplinary collaborative research. Within this case, trust across the three identified dimensions of people, data, and systems were examined inductively to understand the experiences and perceptions of the group (figure 2.5). Aspects such as ability and embeddedness (social and institutional) featured prominently. Those interviewed focused predominantly on their role as data consumers, and thus discussed the trustworthiness cues necessary for them to utilise data and models which they themselves have not created. This included both trust in people and trust in data. Reflective of their roles within their fields, trust in people was mostly interpersonal rather than impersonal. There is an innate familiarity with the field and the process of multidisciplinary and transdisciplinary research that, whilst recognising the difficulties and challenges, these were thought as 'part and parcel' of the work. The interviewees were able to speak to the uncertainty arising

from the lack of direct observation or familiarity occurring during scientific practice, noting that reputation and affiliation were important to ascertain quality and hence trustworthiness, effectively enabling a certain level of confidence by virtue of their membership. Yet, participants were clear to mention that this was not the only trustworthiness cue, for instance as mentioned above Participant 5 noted that data curated by a data centre does not automatically equate to trust in the data by virtue of it belonging in the repository. This speaks to both the literature and the stance taken within this research, that trust is developed using many aspects of scientific and social knowledge to determine of trustworthiness.

As data producers, the interviewees spoke about the challenges they have in presenting and transferring knowledge surrounding their data (section 6.1). When questioned on these challenges and offered potential technological solutions, participants noted that these aspects would still require time. This includes not only the time available in their already busy jobs, but also that a change in funding and grants would need to be created to 'build in' time after projects had finished in order to sufficiently add detail (Participant 13). There was also concern that this could lead to more problems in the future:

> "*If everything's open all the time, it means you never get to move on. So 10 years later, you know, I might have- or 20 years later, I could have somebody contacting me saying, 'oh, I've been looking over this work that you did 20 years ago. And I'm, you know, I'm really interested in it but I think you've made an error or something'. And this is, this is- I have got to go back now and start looking at this thing that I did all this time ago, which, you know, I can neither remember nor really care about if it's such a long time ago. So, how do I always maintain that version control and ensure that the work I did 20 years ago is just as good as the work I'm doing now? That's the real challenge, how do you make reproducible and trustworthy research I think is still something that I don't think has been fully understood yet*" (Participant 4).

This speaks to the issue raised within this research of finding solutions for multiple stakeholders. There is no one-size-fits-all option to addressing issues of trustworthiness in technology design. Despite contact with the original data producer being cited as a key priority in forming trust, it is clear for some that this would be too much for them to undertake and may consequently lead them to either not engaging with the system or not fully documenting their processes.

The creation of trust affordances for this empirical context seeks to represent these nuances. Representing data producers and consumers, rather than curators (section 5.2), allows for a focus on the relevant trust affordances in the same situated context but from different angles, i.e., 'kaleidoscopic patterns' of trust (section 2.5). A greater number of transparency affordances were identified in comparison to provenance and social translucency. The transparency-related trust affordances pertain to understanding in various ways. Scrutability focuses on the opportunity to inspect and scrutinise, whereas abstractability and reciprocality focus on the transfer of knowledge either through the presentation of information in differing forms or through dialogic exchange. It should be noted that through the development of affordances to a specific context, one can begin to see a mosaic appearing. Some affordances interlock and can work towards the same goal for multiple stakeholders, for instance, traceability leads to verifiability identified in both cases. Whereas other affordances illustrate clear advantages *and* disadvantages, which system designers must deliberate before transforming into a system. For instance, awareness and identity are useful for forming trust and facilitating collaboration for trustors but may not be appreciated by those who are the trustees. A comparison of the trust affordances derived from the separate cases will be compared and contrasted in the following chapter.

Chapter 7

Discussion

This research focuses on informing the design of socio-technical systems, such that they are trustworthy-by-design for multiple stakeholders. Specifically, this research has focused on a triad of trust dimensions: people, data, and systems within environmental data science. From the outset, this research has argued for embracing the complexity and nuance of trust when designing a system for multiple – and potentially competing – trust*s* of differing stakeholders. To aid in these efforts, the concept of trust affordances was developed to facilitate these considerations within the design process.

This chapter will appraise the research, including answering the research questions, aims, and objectives and reviewing the research design (section 7.1). In section 7.2, the chapter discusses trust affordances, bringing together the results from the two empirical cases and considering the translation of these trust affordances into the design process. The section finishes by introducing mock prototypes of a Virtual Lab, discussing how features may be instilled based upon trust affordances, and how this process can be used by both system designers and potential system users in order to iteratively develop a final product. The chapter concludes with a wider discussive contribution on this research as a whole (section 7.3).

7.1 Appraising the Research

The overall aim of the research is to inform the design of technology to support well-placed, appropriate trust for multiple stakeholders whilst embracing the nuances and complexities of trust within environmental data science. In order to appraise the research fully, the research questions and their associated aims and objectives (section 7.1.1) and the research design (section 7.1.2) are revisited in order to demonstrate how the research conducted within this dissertation has addressed these questions and in doing so, contributes to the scholarly discourse.

7.1.1 Addressing the Research Questions

The first research question sought to answer *how do stakeholders determine trustworthiness in order to place or refuse trust within environmental data science?* Addressing this aim, *to understand how stakeholders determine trustworthiness*, was attained by defining trust and seeking to understand trustworthiness cues that can be used to determine whether trust should be placed. This was achieved by synthesising related literature in chapter 2, importantly setting the foundation for the remainder of the research. Trust is recognised as a vital yet challenging ingredient within the design of systems (cf. Knowles, Rouncefield, et al., 2015). Whilst we can easily relate to it from everyday experience, defining trust is challenging, often leading to lofty descriptions as *"the chicken soup of social life"* (Uslaner, 2000, p.569) and *"a confusing potpourri"* (Shapiro, 1987, p.625). Social-psychological approaches to trust focus on the perceptions and feelings of trustors, rather than the cognitive examination of information, believing that the presence of perfect information nor the time to consider it are attainable in reality (Blomqvist, 1997; Giddens, 1990; Lewis and Weigert, 1985; Sundqvist, 2011). This approach recognises that trust is on a spectrum ranging from affective to cognitive, dependent on the trustor and the context, impacted by many contextual aspects (section 2.1). The empirical data gathered showed a mixture of processing styles: *heuristic* when a data consumer can

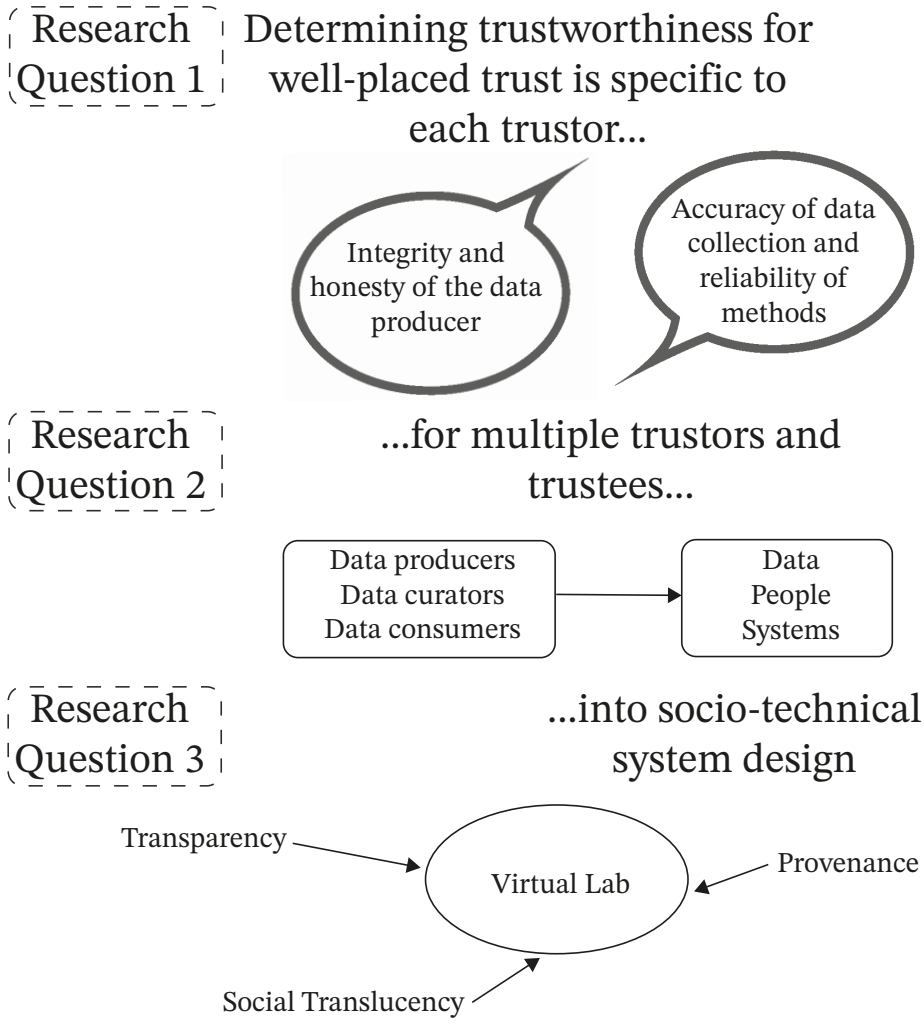


Figure 7.1: An illustrative representation of the three research questions and the key insights found. The image is constructed in a nested style to connote how the research questions operate in a sequential fashion, with the first research question leading into the second question and so forth.

use a rule-of-thumb to assess whether the data fits with their experiential knowledge and *considered* when they need to truly understand the methodology, processing, and presentation of results. Participants also spoke about the context within which data is situated, and the influence of trust in people on their resulting perception of trust in data. A definition was chosen on which to base the research: *"trust concerns a positive expectation regarding the behavior of somebody or something*

in a situation that entails risk to the trusting party" (Marsh and Dibben, 2003, p.470). The definition, being suited to a general context involving both people and systems, is suitable for the research at hand and does not contradict the empirical findings generated. Risk and uncertainty are a general feature of scientific research limited to not only the data one uses but the people one collaborates with. There was evidence of positive expectations, but following the application of two models of trust a definition including a descriptor other than 'behaviour' might be more appropriate, given that this applies only to trust in people and trust in systems, but not of data (when data is the trustee). Data does not behave one way or another, and thus a more appropriate wording might be 'character' connoting the essential qualities of either 'somebody or something'.

The second part of the research question focuses on trustworthiness cues. It is known - but perhaps not always appreciated - that trust can be inferred from multiple cues of trustworthiness by a single individual and that these cues will vary between people. The rationale behind this research question was that rather than a system tell a user that something is trustworthy by generating a metric or score, it is much more beneficial to cater to the prevalent trustworthiness cues within a context and allow them the opportunity to align perceptions of trustworthiness and well-placed trust. As discussed within chapter 2, cues of trustworthiness vary. For instance, Zimmerman (2008a) found that ecologists use multiple data sources frequently and are adept - from a combination of disciplinary training and experience - at determining whether to trust data. Their ability to understand data was most important for them in reusing data, rather than knowing personally, or by reputation, data producers (Zimmerman, 2008a). Achieving this level of understanding due to disconnection between data producers and consumers is challenging, raising "the central question as to how data collected or constructed by one researcher can be trusted or even understood by another" (Carlson and Anderson, 2007, p.643). To this end, the chapter sought to understand both technical and social aspects of trust in data (Birnholtz and Bietz, 2003), culminating

in a developed understanding of how purported trust model components speak to different cues of trustworthiness, e.g., reliability, functionality, or forms of embeddedness (figures 2.4 and 2.5). Regarding the findings from the empirical data in chapters 5 and 6, attaining a level of understanding was important for the participants interviewed. This is considered an essential part of their ability to reuse data. This runs contrary to how trust is thought to operate in other contexts, i.e., as a mechanism to reduce the costs of monitoring subordinates (Cook and Santana, 2020; Luhmann, 2017). As noted in section 2.6, it is assumed that information may not be checked if a trustor already trusts a trustee, but this is antithetical to the reality of scientific pursuits. The participating environmental data scientists highlighted the importance of validating data regardless of the data producer, when that data was essential to their endeavours. They did not claim that this was because of a lack of trust in all instances, but did note that data often comes from varied sources and thus varied producers. Simply, it is about vetting data quality for use. When they knew the producer this helped them to form trust, but scrutiny still remained an important step in their process. They might be more or less trusting prior to this exploration of quality, e.g., if they have familiarity with the data producer or can see a higher level of detailed information included with the dataset, which lead them to scrutinise to greater or lesser extents, but this stage occurs regardless of trust – assessing data for appropriateness, quality, and completeness is a standard scientific starting point (Douthit et al., 2021; Yoon, 2017). This method is more characteristic of science rather than other forms of information, given its high standards, potential impacts, and vulnerability of scientists in producing potentially incorrect results. Regarding trust in data, authenticity, accuracy and credibility of data and the identity of data producers are important characteristics in determining whether to place trust (section 2.2). Owing to the potential sparsity of appropriate data in some sub-disciplines, some participants were pragmatic about quality concerns but a common theme was that being able to determine quality and understand the context including assumptions and biases surrounding data through

access and visibility was essential in order to utilise and trust data. This is different from a general understanding such as being able to understand research published in a paper that is not of direct relevance to the research at hand or being told a fact given in a conversation, which might be more dependent on trust in the trustee than data used in science. Thus cues of trustworthiness may also be different within science to other social arenas. The empirical chapters found that a whole range of cues were relevant to a range of stakeholders. Provenance and contextual information (including meta-data and additional supplementary documentation) were highlighted across both cases. But appealing to these means would be different for a sub-disciplinary expert in comparison to a trustor who is not as familiar in that specific field.

The second research question asks *what are the salient types of trust within environmental data science and how do they differ?* Within environmental data science, trustees (those in which trust may be placed) and trustors (those placing trust) take multiple forms. Stakeholders were divided into three groups: data producers, data curators, and data consumers. At the outset of the research, the necessity of balancing these different perspectives was highlighted (chapter 1), alongside an introduction to the multiple trustees present. Trustees include: data, people (all three categories above), and systems all of which are interrelated and interlinked (Jantz and Giarlo, 2007; Knowles, Rouncefield, et al., 2015; Knowles, 2016; Marsh, Atele-Williams, et al., 2020). Focusing on any aspect, i.e., environmental data, must be made in conjunction with the intersecting trustees in order to obtain a complete picture. Failing to attend to these aspects within research will return an incomplete picture of reality and thus a flawed end design. Chapter 2 sought to address the aim to conceptualise trust within environmental data science. This was achieved by analysing two prevalent and well-cited trust models within the literature. Trust can be interpersonal (between people); interpersonal but mediated through a system (e.g., co-located work, wherein trust is placed in both the person and the system); or between a person and a technology, where the

system is trusted. Virtual Labs can potentially be the objects of trust themselves (trustee); contain data and models which can also be objects of trust; and can, when designed for, enable communication between users (mediator) as shown in figure 2.2. Because there are three differing dimensions, no singular model exists to represent the context. The mediator model focused on people and could pertain to socio-technical systems whereas the trustee model focused on systems. The research addressed the aim to *conceptualise the nature of trust with respect to three dimensions of trust*, both conceptually and latterly, empirically. In chapters 5 and 6, trust in data is formed directly in data as a trustee and mediated via trust in people. When interpersonal trust did not exist, impersonal trust, i.e., trust in institutions, can often be a replacement for interpersonal trust (Cook and Santana, 2020). The institutions in this instance are the research and data centres run by data curators (chapter 5). For data curators, they viewed this institutional trust as salient for both data producers and consumers. This is expected given their roles and potential biases, it would be concerning if they *did not* view institutional trust as important. But for those scientists studied within the Centre of Excellence case, institutional trust did not feature highly. This is a potentially implicit aspect of trust that is not often considered given their expertise. This form of trust is not as important to them in comparison to other less-experienced environmental data scientists or researchers with different disciplinary backgrounds, i.e., they have a wealth of experiential knowledge and personal connections that make impersonal trust a background element to their placement of trust. If these sources of trust were not present or if trust in institutions had in the past been brought into question, then one might anticipate a higher consideration of this. Regarding trust in systems, distributed ledger technology and Virtual Labs were used to explore the potential for technological development. Reflecting on the empirical insights gleaned, technology can work to foster well-placed trust – if intentionally designed to – by appealing to trust in the technology itself, alongside trust in people and trust in data. For instance, ability was a key component of both trust in people

and trust in systems, speaking to expertise, predictability, and competency. Forms of embeddedness like internalised norms, were not often explicitly cited as a reason to trust. Internalised norms were often not explicitly stated but were perceivable when analysing the empirical data, inferred in both people and systems, related to honesty, credibility, and consistency. Benevolence did not feature prominently for either data curators or those who are both producers and consumers other than when discussing the time and effort involved on their behalf in uploading and describing data and associated information. Impersonal forms of trust – which can come into play when trusting in people either by situating the trust (reputation) or circumventing a lack of interpersonal trust – are often implicit and subconscious in real life, making extracting these insights challenging. Overall, regarding the dimensions of trust, trust in people and data were the most prominent with ability being identified as important in each dimension, followed by the three forms of embeddedness (institutional, temporal, and social). Trust in systems was relevant as a mediator, but less relevant as a trustee.

The third research question builds upon the preceding questions, seeking to answer *how can considerations of trust be incorporated into system design?* Owing to the complex nature of trust, the formalisation of the findings into system design is non-trivial. To address this gap in knowledge, and to account for complexities and nuances, the third aim of the research was to *develop the concept of trust affordances to foster well-placed trust within Virtual Labs.* Trust affordances were developed as a means to attend to multiple trustees and the ways in which they place trust. Following the work undertaken in chapter 2 to understand the utility of trust models, the theoretical concept of trust affordances was developed (chapter 3) to support the overall research aim. Trust affordances are defined as "characteristics of the technology by virtue of itself or of features designed into the technology to promote trust by providing access to evidence of (dis)trustworthiness specific to a user, a technology, and their context" (Thornton, Knowles, and Blair, 2021, p.67). Rather than models which can be static, trust affordances were developed to create

identifiable instances in which trust could be fostered whilst speaking to multiple stakeholders. These were organised into three categories (provenance, transparency, and social translucency) that are often connected as antecedents to trust within the social computing literature (Thornton, Knowles, and Blair, 2021). The identified trust affordances speak across and between model components (i.e., those characteristics featured in figure 2.5). For instance, verifiability (provenance) and visibility (social translucency) both speak to the ability to ascertain what is accurate and true for both data and people respectively. Similarly, comprehensibility (social translucency) and scrutability (transparency) speak to the ability to understand in helping to form trust. Trust affordances were identified within each empirical chapter (sections 5.2 and 6.2).Trust affordances can be valuable to system designers when considering trust and present a means to both explore contexts and develop understanding prior to instantiation, discussed further in sections 7.2 and 8.2.

7.1.2 Assessing the Research Design

The research design speaks to all three research questions, with objectives to identify and conduct an appropriate research design, to conduct qualitative empirical research, and to analyse and re-analyse the data gathered (figure 1.2). A case study approach was chosen as this aligned with the constructivist paradigmatic stance and social-psychological approach to trust taken within the research (Appendix A.2). This provided a base to seek a variety of participants to listen to multiple voices (Mertens, 2014) with qualitative, interpretivist methods chosen as the ways to address the research questions and to fulfill the research design.

The research approach was crafted to address the research questions and aim of the research. The purpose of both case study and qualitative research is not to test hypotheses, measure variation, or to generalise to populations, rather it is interpret contexts and participants' meanings (Bryman, 2008; Maxwell and Mittapalli, 2008; Maxwell, 2009; Sandelowski, 1996; Yin, 2017). The findings of interpretivist inquiry are always conditional, context-specific, and contingent (Angen, 2000; Charmaz,

2009; Mishler, 1990). This aligns with social-psychological approach to trust, viewing trust as situational between a trustor and trustee in a specific instance (Cofta, 2021; Marsh, Atele-Williams, et al., 2020). This makes the evaluation of such research challenging. Constructivism and natural forms of inquiry focus on valuating research in terms of trustworthiness rather than as having validity or being 'the truth' (Creswell and Miller, 2000; Lincoln and Guba, 1985; Mishler, 1990). "Reality is multidimensional, deep, and complex and [...] understanding is necessarily partial" (Davis, 2008, p.756). Standard forms of evaluation are not suitable because qualitative research methods exhibit such breadth that they cannot represented by any singular set of criteria, rather, the quality of each research project should be determined independently (Lincoln and Guba, 1985; Lub, 2015; Sandelowski and Barroso, 2002; Tracy, 2010). Angen (2000) states that conclusions are open to interpretation, but the researcher must demonstrate how they came to their interpretation and why it is worthy of trust by demonstrating thoughtfulness, usefulness, respect of diversity, and openness (e.g., taking the conversation in new directions and creating new questions). Mishler (1990) suggests asking: "what are the warrants for my claims? Could other investigators make a reasonable judgement of their adequacy? Would they be able to determine how my findings and interpretations were produced, and, on that basis, decide whether they were trustworthy enough to be relied upon for their own work?" (p.429). To be accepted as trustworthy, researchers must demonstrate their consistency, dependability, systematic analysis, and must disclose in enough detail to allow for evaluation (Costantino, 2008; Nowell et al., 2017).

In order to circumnavigate this complexity, the research has been written up thoroughly to enable evaluation and additionally, a number of methods of evaluation were included within the research design. This includes multiple data sources and an openness to interpreting the views of participants as they speak them (Tracy, 2010). A heterogeneous and small sample was selected in each case, characteristic of a case-based approach (Roy et al., 2015; Sandelowski, 1996). This is also

reflective of the 'fragmented' nature of environmental data science (Blair, Bassett, et al., 2021), thus in both empirical cases, participants were sought from different environmental data science sub-disciplines in order to capture a wide array of views. The data was assessed for 'informational redundancy' (data saturation) whereby no new information is garnered from additional sampling (Fusch and Ness, 2015; Lincoln and Guba, 1985; Mason, 2010). Fusch and Ness (2015) suggest that rich and thick data is a useful marker for determining saturation: "thick data is a lot of data; rich data is many-layered, intricate, detailed, nuanced" (p.1409). Throughout this research, quotes have been included to contribute to both quality (richness) and quantity (thickness) (Denzin, 1989; Fusch and Ness, 2015), allowing readers to gain understanding and to therefore judge the transferability of findings to other contexts (Braun and Clarke, 2006; Lincoln and Guba, 1985; Nowell et al., 2017). Reliability can be ascertained through the consistent patterns of themes (Creswell, 2013) and attempts to present the data to multiple audiences (Bryman, 2008; Stake, 1995). The workshop and survey were designed to fulfil these evaluations, by evidencing that no new information was garnered from additional sampling (Fusch and Ness, 2015; Lincoln and Guba, 1985; Mason, 2010). Additionally, literature that (dis)confirmed the data collected was included to assess the validity of the research.

In addition to the above, "qualitative researchers should be as concerned to produce a moral or ethical research design as we are to produce an intellectually coherent and compelling one" (Mason, 2002, p. 41). Creswell (2013) and Tracy (2010) also agree that high quality qualitative research is that which is ethical. This includes seeking approval from relevant ethical committees but there are also further considerations such as relationships with participants and power imbalances alongside fair and equal reporting of all perspectives (Creswell, 2013; Mason, 2002; Tracy, 2010). Researchers must also seek to report all perspectives given, not just those that align with their own beliefs and values and should additionally report findings in a clear way for the intended audience(s) (Creswell, 2013). The ethics of this research have been undertaken seriously, with participants' consent sought in

every instance, along with the following of detailed ethics plans and a presentation of participants' views as expressed. Finally, researchers must also consider the impacts of this reporting, e.g., on funding bodies and participants (Mason, 2002). This research is not funded by the Research Council explored within chapter 5 and every effort has been made to ensure that no participant is identifiable to their quotes or views expressed to promote their confidentiality.

Overall, the research design adopted has been well-suited to addressing the research questions. The methods chosen allowed for a thorough exploration of the phenomenon of trust within two original cases. Rather than quantitative measures, these methods were chosen to gain perspectives and insights from a range of participants. The flexibility of qualitative research was also important, allowing for the inclusion of data re-analysis at a later date with respect to trust affordances (figure 4.3). Potential limitations to this approach include the notion that multiple different methods and frameworks are not ways of finding a valid singular truth but open up the complexities of the issue (Tracy, 2010). Every effort was sought to address the research questions with multiple, yet appropriate, methods. As is the case with qualitative research one might always argue that there is never 'enough'. Rather than include more methods to the existing cases, a potential avenue might be to include different stakeholders, for instance segmenting data producers and consumers further by targeting different demographics. Another limitation might be the ability to generalise further, a common criticism of case study approaches and of qualitative research (Bryman, 2008; Priya, 2021). However, the goal of this research is provide trust affordances to aid those in designing a trustworthy socio-technical system, rather than to produce a specification of the design of technology in a specific instance within a specific scientific domain. Trust affordances are designed as a concept to be transferable and reusable in multiple contexts. And, importantly, a point which this research has made great effort to state throughout: neither the findings of the empirical research nor the trust affordances developed from them should be generalised to other contexts, they are specific to both a time and a place.

7.2 Trust Affordances

Trust affordances were developed as a concept to deliver practicable insights into designing for trust, building upon the finding that models of trust are useful at simplifying a complex concept but less useful at reflecting the complex reality of trust, ultimately impacting upon translation into design. Models of trust are static and segment the components of trust into diagrams comprised of neat boxes (section 2.6). We can either generate general models that have a wider application, or we can take time to understand the context and the trustors and trustees within that context, in the hope to generate a model which can account for that intersection, but which may lack the generalisability of broader models. This research contends that given that what 'A trusts B to do X' is so inherently variable based on the multiple combinations of A, B and X, no model can ever hope to account for this adequately. Moreover, in precisely what ways A feels able to trust B may not be segmented into neat components for A in their experience of trust (Jones and George, 1998). To this end, trust affordances are as a means to gain further insight prior to design. The rationale behind trust affordances is the inherent alchemy required to develop systems that cater to multiple stakeholders' trust and "the need for designers to engage with the richness of the trust landscape" (Thornton, Knowles, and Blair, 2022, p.1387). Trust is complex and especially so in the context of this research: to inform the design of a socio-technical system that accounts for multiple stakeholders in multiple contexts with multiple trusts. Like trust, trust affordances are conceptualised as relational, so may or may not appear as an opportunity for a specific user dependent their perception.

7.2.1 Comparing Trust Affordances Between Cases

The two cases represent different perspectives of the same context from the perspectives of data producers, data curators, and data consumers. Across the cases, nine trust affordances were identified, presented in table 7.1.

Category	**Case 1: Data Curators**	**Case 2: Data Producers and Consumers**
Provenance	Traceability	Traceability
	Verifiability	Verifiability
	Immutability	
Transparency	Scrutability	Scrutability
		Abstractability
		Reciprocality
Social Translucency	Visibility	Perceptibility
		Comprehensibility

Table 7.1: Identified trust affordances and their related category identified across both empirical research cases.

With respect to the research conducted across both cases, all three categories of trust affordances could be identified with provenance being the highest category (three trust affordances) in the first case, and transparency being the highest (three trust affordances) in the second case. Social translucency was not as strongly represented, and was related primarily to the intricacies of distributed collaborative work, particularly the visible evidence of real-time working and of understanding discipline-driven vocabulary differences.

Some trust affordances align neatly, speaking to the same trust-related aim. For instance, verifiability, scrutability, and comprehensibility focus on the ability to understand through examination and to prove the authenticity of data and results. Similarly, traceability and visibility speak to identity and its importance in determining reputation. Reflecting on the dimensions of trust studied within this research: trust in people, data, and systems, trust affordances predominantly relate to trust in people and data mediated by the technological system. In this regard, systems may or may not afford trust as a trustee directly, but rather through the

features and functionalities they provide. This aligns with the findings in chapter 2 wherein the mediator role trust model was found to be more applicable than the trustee role trust model (figure 2.4). A system will be trusted in some respects, albeit implicitly, i.e., to provide secure communication, to represent information correctly, and to draw together the correct datasets requested.

Of the identified trust affordances, there was limited support for some, for instance immutability could be perceived as both a positive and negative trust affordance, and likewise, abstractability could be perceived as a trust affordance for other stakeholders with a consequential impact on data producers to provide this information. Different perspectives from different stakeholder groups means that varying perspectives have been gleaned on trust within environmental data science. Trust affordances are not applicable in *all* instances within the same context, suggesting that the tensions between stakeholders will prevail unless these differences can be accounted for prior to instantiation. The construction of trust affordances as a concept specifically drew upon the context-dependent nature of affordances and their subjective perception by system users (section 3.3.1). System designers can aim to create trust affordances, i.e., the opportunities to foster well-placed trust, but cannot guarantee their identification and interpretation.

7.2.2 Considering Trust Affordances and Virtual Labs

The research focuses on the creation of trust affordances to inform the design of technology to support well-placed, appropriate trust for multiple stakeholders whilst embracing the nuances and complexities of trust within environmental data science. Virtual Labs were chosen as a technology to facilitate trust in this instance. This is not restricted to the specific instance of the technology, but rather inspired by Sellen and Harper (2001), the use of a technology can bring to the fore characteristics that may be otherwise invisible. For this reason, distributed ledger technology was considered within the first case and the introduction of Virtual Labs over standard data repositories was considered in the second case. As illustrated in figure 1.1, these

infrastructures can be used along the data science pipeline to conduct research, to disseminate results, and to collaborate with others. Whilst not a panacea to issues of trust, their flexibility allows for multiple features to be included during design and development dependent on the context and user's needs, Virtual Labs can either share knowledge as artefacts or support communication access to those with expertise (Ackerman et al., 2013), presenting a suitable option for affording trust over and above solely introducing distributed ledger technology, for instance.

As affordances are user-dependent, a system designer can, following the exercise of categorising potential trust affordances, begin to consider the features and functionalities that may (or may not) lead to fostering well-placed trust for users. As argued throughout this research, trust "may need to be addressed in a multi-pronged, trust-building strategy" (McKnight, Choudhury, and Kacmar, 2002, p.21). For data usage, this includes identity and affiliation of producer (in order to ascertain ability and quality); sufficient meta-data and supplementary information (to determine accuracy, reliability, and appropriateness); and the opportunity to scrutinise the provenance of the data amongst various aspects including lineage, context, and aggregation and processing. These findings align with the literature, for instance the opportunity to inspect data within a system has previously been identified as a necessity within environmental data science (Henrys and Jarvis, 2019; Meyer, Weigelt, and Kreft, 2016; Pescott, Humphrey, and Walker, 2018). Similarly, the notion that data reuse and assessments of trustworthiness involves multiple cues has previously been put forth (Faniel and Yakel, 2017; Wallis, Borgman, et al., 2007; Yoon, 2017). Throughout both empirical cases, participants were asked if they thought any specific elements of the technologies would contribute to the affordance of trust. These are not suggested as recommendations, but can be useful for future system designers to frame their enquiries (Voss et al., 2009). A tension between permanence and additionality was identified in the data curators case, wherein data should exist for the long-term in it's original state, but should also be able to be annotated, linked to, and enhanced in order to become a 'living document'.

This aspect can, if designed appropriately, provide visibility, accessibility, and comprehensibility to other users, enabling them to infer thinking behind analyses and identification of other users' work (Jirotka et al., 2005; Martin, Rodden, et al., 2001). Given that identity in the form of name, affiliation, and ability to contact were identified by participants in both empirical studies and within the survey, efforts to include these elements within a Virtual Lab may be beneficial to a data consumer. Hyperlinks and photographs can be used to aid identification and promote seamless use (Artz and Gil, 2007; Riegelsberger, Sasse, and McCarthy, 2003a). To aid trust through abstractability, narrative can be added to analyses via interfaces such as RShiny[1] and Jupyter notebooks[2] to add interactivity and aid exploration (Bates, Lin, and Goodale, 2016; Dourish and Cruz, 2018; Hollaway et al., 2020). For data producers, this may aid in their lack of trust that their data will be taken and used inappropriately, but as discussed, an ongoing concern for data producers is the amount of epistemic labour that these additional forms of information will take to produce. This concern was also raised with respect to comment boards and direct messaging. Despite the trust afforded for data consumers in these forms of communication, e.g., reciprocality and scrutability, data producers questioned the level of effort this would require of them presently and in the future.

One can see that there are multiple options for affording trust within Virtual Labs, this would involve rigorous and iterative user-testing with multiple sets of stakeholders prior to instantiation in order to determine the extent to which Virtual Labs would afford or preclude trust for users within the context. A common, and as of yet unresolved, challenge is the design of a system to satisfy multiple stakeholders. For each trust affordance and for each function or feature, one can see both advantages and disadvantages that this would bring about. Ultimately, one might argue that design decisions will have to be made and evidence for these decisions will also have to be communicated. Addressing grand challenges facing the planet and people is not advanced by lone inventors holed up in laboratories, it

[1] `https://www.rstudio.com/products/shiny/` accessed February 2024.

[2] `https://jupyter.org` accessed February 2024.

is advanced by a heterogeneous knowledge-seeking community (Oreskes, 2019). In the quest for trust in science at a higher-level, these decisions may have to privilege the trust of some stakeholders over others, dependent upon the context and at the very least, a careful exploration of the multi-sided nature of trust affordances is an undeniable necessity for anyone involved in the design and development of socio-technical systems.

7.2.3 Translating Trust Affordances into Design

To demonstrate what this research means in practice for system designers, this section presents illustrative prototypes of Virtual Labs. In what follows, instances of features are presented and explained with respect to the relevant trust affordances. This is beneficial not only to designers themselves, but may also be included as a means to communicate design decisions to technology users (Gaver and Bowers, 2012). As discussed above (section 7.2.2), there are many identifiable features of a Virtual Lab that may afford trust for different stakeholders. Thus far, the research has sought to remain at a generative, intermediate-level (Löwgren, 2013), such that specific instances of design solutions are not given, to avoid the perception of prescription ("here's what works for any and every context"). This section therefore illustrates what or how these trust affordances may appear within design, when perceived by users as affordances for opportunity.

Figure 7.2 shows three mock prototypes of Virtual Lab windows, created to show how screens might appear prior to user interface design during a co-design process, with iterative prototypes developed following consultation and feedback. The top pane is representative of a homepage for a dataset, the middle pane shows the file contents for the dataset, and the bottom pane shows an example of one of the datasets: data(2).csv with 'User X' logged into all three windows. The mock protoypes are interlinked, based on 'Lorem Ipsum' dataset where the second pane shows the contents of a folder shown in the first pane and the third pane shows a file contained within the folder on the second pane.

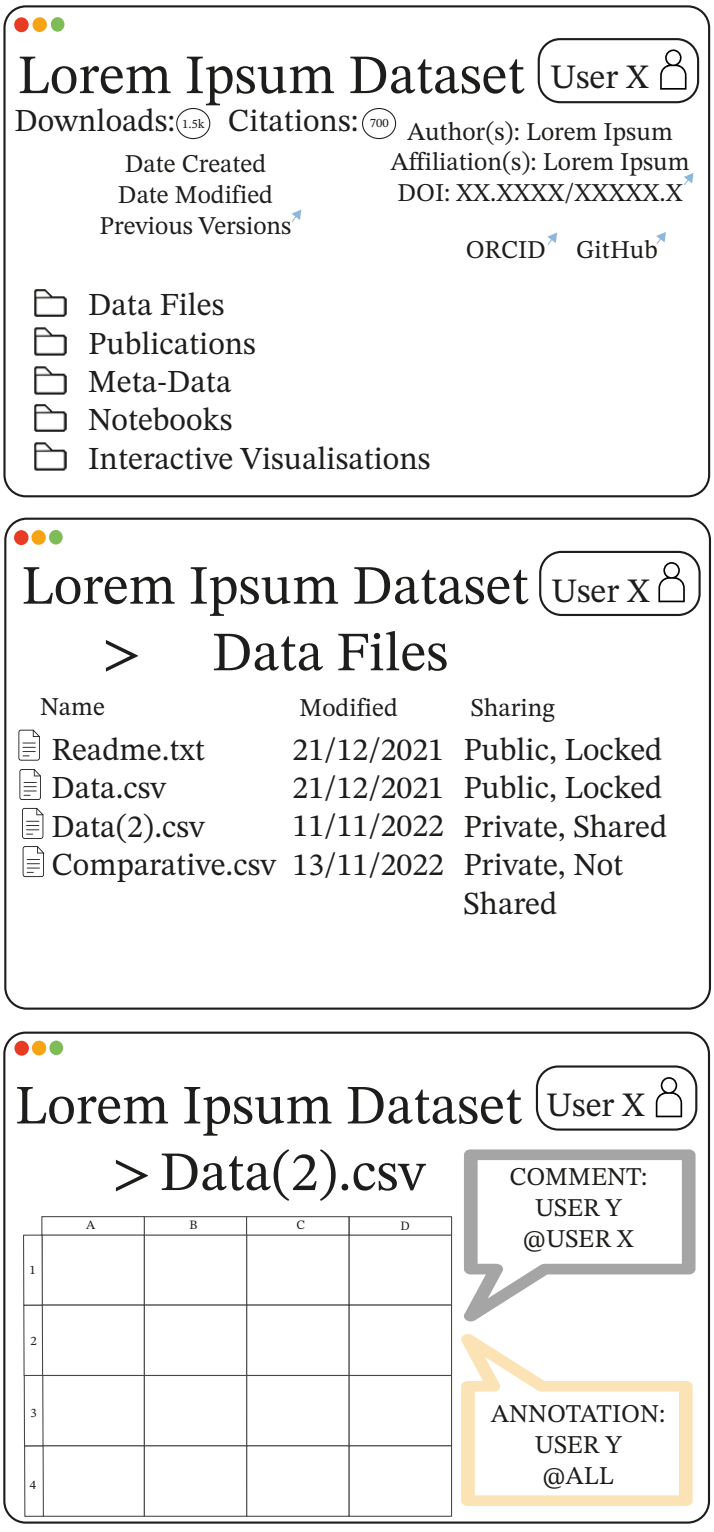


Figure 7.2: Mock prototypes of a Virtual Lab, showing three different screens: homepage, file contents, and a dataset. Blue arrows indicate hyperlinks.

The first pane (top of figure 7.2) includes information about the authors and their affiliations; DOI for the data; hyperlinks to ORCID[3] and GitHub[4] accounts; downloads and citations for the dataset; a hyperlink to previous versions of the data; and files related to the dataset. In terms of trust affordances, 'visibility' (social translucency, section 5.2.3) may be afforded with the inclusion of authors and their affiliations, providing easily accessible information about data producers. External hyperlinks such as DOI, ORCID, and GitHub also contribute towards this, whereas names might be useful for reputational assessment if data producers are known, links to their other work also seeks to improve visibility. These external hyperlinks also link to different versions of the work, i.e., to an academic paper or to the original code, thus speaking to elements of 'abstractability' and 'scrutablity' (transparency, section 6.2.1). Downloads and citations may act as a means to determine trustworthiness, rather than a trust metric or score which may or may not personally relate to a user, these figures can act as means for a data consumer to determine whether more citations is equal to higher trust. If these features are designed to link to datasets, models, or papers which have used this data an element of 'traceability' may be afforded (provenance, sections 5.2.1 and 6.2.2).

The second pane (middle of figure 7.2) focuses on the same dataset as above and shows the contents of the 'Data Files' folder. The most relevant feature within this pane is the 'sharing' column, providing information as to whom this file is accessible by. 'User X' is logged into all three screens, but now their presence shows that two files are private, one is shared and one is not ('data(2).csv' and 'comparative.csv'). Being able to identify which spaces are secure, i.e., who has access, was raised in relation to social translucency ('perceptibility' in section 6.2.3). By simply showing which are private, public, or shared with select users, this feature directs a user to both their audience in two senses: who can see their actions and to whom they are speaking.

The third pane (bottom of figure 7.2) shows a csv file, and it is at this point

[3] `https://orcid.org` accessed November 2024.

[4] `https://github.com` accessed November 2024.

that we can see that because 'data(2).csv' is private but shared, there is a comment from 'User Y' to 'User X' which is private and an annotation to the dataset which is public. This pane shows that both direct comments and public annotations are available functions with respect to a dataset, relating to the suggestions made for the transparency-related trust affordance of 'reciprocality' (section 6.2.1). This also brings awareness in the form of 'perceptibility' when collaborating with other users (social translucency, section 6.2.3) and touches upon the tensions of 'immutability' and the ability for living documents (provenance, section 5.2.1).

Whilst consisting of a low-fidelty in terms of design aesthetic, these mock prototypes can be used within the design process to communicate between system designers and stakeholders if co-design is being undertaken. These mock prototypes can also be used to provide a narrative of the process of system design when the Virtual Lab is built. One can imagine a communicative process of how and why certain features were designed into the system. These features are not meant to suggest that any and all features are constitutive of a set of trust affordances, rather to demonstrate the connections between the two. Trust affordances can be organised in a similar way to models of trust (see for instance Thornton, Knowles, and Blair (2021, figure 1, p.71)) and can be used in conjunction with models of trust if system designers so wish. However, unlike conceptual trust models, trust affordances recognise that each element of trust can exist on a spectrum (rather than being a neat box, an element might instead vary in terms of opacity for each person) and that these elements can overlap (instead of being neatly disparate components). Fitting trust models onto a particular context is challenging (chapter 2). Rather than producing a model of affordances, or a hybrid of a trust model and affordances, this section illustrates how affordances can be identified within specific system features.

Of course, not all trust affordances can be represented within three windows, but during the co-design process one might opt to focus on specific trust affordance categories, e.g., in what forms provenance is best designed into a Virtual Lab: *is a visual map of a data journey comparable to a thorough context-rich supplementary*

document? How might we introduce verification of this provenance information? Instead, one might choose to focus in fine detail on each individual feature. For instance, the presence of a DOI for the Lorem Ipsum dataset and a folder containing meta-data in the first pane allude to the possibility for 'immutability' and 'verifiability' (provenance, section 5.2.1), yet how these are specifically designed for will impact upon the perception of users. Meta-data can be extended variously, dependent upon the provenance information provided by data producers and collected by data curators. The inclusion in figure 7.2 is intended as illustrative of the principle, as covered within this research, the quality and format of this meta-data will need to be thoroughly examined alongside the inclusion of user-generated content and other methods of creating context (section 6.3). The example within figure 7.2 also does not include the specific system features for transparecy-related trust affordances such as 'scrutability' (sections 5.2.2 and 6.2.1) or 'abstractability' (section 6.2.1). These affordances are indirectly referred to with the presence of datasets, notebooks, and interactive visualisations, yet will ultimately depend on the final configuration determined by stakeholder need. Part of this design process may also choose to consider: *what makes different iterations more trustable to different stakeholders and why? How do these prototypes complement or improve the cues available to them currently for fostering well-placed trust?*

This section shows how trust affordances identified within the empirical data can be used to contribute to system design, and importantly, provide a demonstration of design decisions. The inclusion of an excerpt of these prototypes illustrates that individual windows can be focused upon to show different trust affordances and one can imagine that a whole schematic consisting of many windows of a system, with multiple panes allowing for insights into the design process. In terms of the design process, the research approach outlined within this dissertation may be replicated by others. First, one must attain an understanding of stakeholders and context, this promotes an insight into the multiple aspects of trust present in a setting. This provides the foundation for conceptualising trust affordances, which may then

be refined through ongoing reflections with stakeholders before instantiation into design. As a by-product of this process, these efforts can be used to communicate with future system users about the design choices made. Trust affordances should be derived from the stakeholders themselves rather than pre-determined by system designers. Trust affordances, whilst not named or identified as such by authors, can be drawn from related ideas within the literature and utilised as inspiration. The act of defining affordances is akin to a liminal space found between the transition from the reality of a stakeholder, their perceptions and needs into a representation of a system and instantiation into design. Although one can follow the process laid out within this work, the art (thinking beyond trust types and categories, system features and functions) and alchemy (transformation and creation of existing constructs and theories) should not be discounted as a vital ingredient (Thornton, Knowles, and Blair, 2022).

7.3 Further Remarks

Given the insights discussed thus far, there are some final thoughts that are worthy of inclusion, namely: a consideration of the context, identifying the primary use of technology, and the inclusion of a social infrastructure within technological design.

This research was conducted within a specific context. The study of data curators in chapter 5 was perfectly, albeit coincidentally, timed in that both the institution and those interviewed were undergoing a period of transformation. In this respect participants were primed and receptive to the introduction of technology. All but one (Participant C, 'Blue Data Centre') were keen to discuss the nature of trust and potential impacts. Given this context, the Research Council were cognizant of their role as a trustee and were willing to demonstrate this to a variety of stakeholders through different means both in the interviews and in the following workshop. In a different context the institution may not be so open to either discussing trust issues or being open to the possibility of technological introduction.

Anecdotally, in conversations with both research centre and data centre employees and with environmental data scientists following the research presented within this dissertation, the thoughts on trust and future technology design are the same as those reported, though one must be careful to extrapolate this finding beyond this context, i.e., to another scientific discipline.

There are two particular challenges facing trustworthy system design within environmental data science: distributed virtual teamwork and working with unfamiliar data, models and people (chapter 1). Both of these challenges were apparent within the research, but most participants spoke to the challenge of working with data-related unfamiliarity than of virtual teamwork. In most instances, participants had enough experience that they either knew personally or knew of their collaborators for project-work, but this was more challenging when selecting data. Designing systems for known groups of users *that also know each other* introduces limits and boundaries that designers can work with and assumptions to be made. Trust in people as a dimension includes individuals, social groups, with impersonal trust in institutions as a supporting factor. Thus a potential inclusion for system designers utilising this research could be an expansion of the trustors and trustees that they consider dependent on the primary use the technology is likely to take.

Finally, a point that should not be overlooked in the design of systems is the role of humans, as advisors and administrators, to support its ongoing functionality (Baker and Yarmey, 2009; Martin and Rouncefield, 2003; Møller et al., 2020; Wadlow, 2014). A socio-technical system is not just the system itself, but the infrastructure surrounding it. Retaining a human in the loop is potentially a long-term cost greater than the initial costs of design, development, and deployment but should not be ignored. To foster trust in the future, those responsible for systems should seek to commit to ensuring the role of people. This is increasingly important when the introduction of a new system is undertaken, funding might be provided for a certain period of time, and thus those developing a system should take care to ensure that there are provisions for the future. Not doing so will likely impede the

opportunity for users to build trust in the system or to engage with the system at all. Clear communication on this will be a source of trustworthiness for some users.

Chapter 8

Conclusion

In the quest to address pressing challenges facing both the Earth, and its people, plants, and animals, environmental data science seeks to bring together transdisciplinary groups. Trust is an important facilitator of collaboration and core component of the resulting outcomes. Yet designing a system to improve distributed virtual teamwork and to aid the use of working with unfamiliar data, models, and people is challenging. Thus, this research sought to uncover the myriad ways in which trust is perceived in order to address these difficulties. Trust "contains multitudes" (Thornton, Knowles, and Blair, 2022, p.1393) and this research has, at every stage, sought to embrace the complexities and nuances, rather than to simplify and shortcut how we, as system designers, may facilitate trust or demonstrate trustworthiness. This final chapter reviews the contributions of the research (section 8.1); cements the importance of trust (section 8.2); and sets out the agenda for future work (section 8.3).

8.1 Review of Contributions

This research makes core contributions to the overlapping research areas surrounding social computing, trust, and the study of scientific practice through synthesis of a wide array of literature. This includes an advancement of the understanding of trust

within environmental data science. Despite this being a strategic theme set out by the Natural Environment Research Council (2022), to date no work had achieved the comprehensive – and importantly, nuanced – understanding of trust needed to enact their programme. This work has contributed an important insight into this domain, providing both a detailed description of trust across multiple stakeholder groups and dimensions. This will be of benefit to the Research Council with their quest to facilitate confidence and trust in digital research. Beyond a focus on environmental data science, this research is valuable to the social computing domain, particularly CSCW (Computer Supported Cooperative Work), wherein trust is a recognised ingredient of collaborative work, but remains an ongoing, and elusive, goal. This research has further illuminated trust both in and of systems. The goal of this research at the outset was not to produce a schematic diagrammatic representation of a Virtual Lab, but to inform a new sensibility for approaching system development, in line with a trustworthy-by-design approach (Knowles, Harding, et al., 2014). The output of this research will be directly relevant to those currently working towards the design of trustworthy systems across a broad array of contexts beyond environmental data science, in any context in which trust is needed "not only in the scientific findings and resulting recommendations, but subsumed within this, trust in the people, processes, and data that led to these findings and recommendations" (Thornton, Knowles, and Blair, 2021, p.64).

8.2 The Importance of Trust

Data producers, curators, and consumers alike value trust within environmental data science in order to produce trustworthy results and advancements in the area. For data consumers, being able to trust the elements utilised within their work helps them to create trustworthy research. With awareness of the particularities of data they use, they can ensure that they are using the most appropriate data for their analyses and can calculate and describe the varying types of uncertainties

within their outputs. Given that reputation was a large component of trust in people, relatively few participants described how they themselves work towards being perceived as trustworthy by others and by what means. For data producers, the most important trust issue for them was trust in data consumers, that their data would be taken and used appropriately. This is increasingly prevalent when those using one's data cannot be assumed to be of the same community of practice, impacting upon what can be taken for granted: similar methods and standards when collecting, analysing, and interpreting data; speaking the same scientific language; and similar paradigmatic stances including ontology and epistemology (Van House, Butler, and Schiff, 1998). One might suppose that this would lead them to document their data sufficiently for a wide audience to facilitate this process, but in reality a repeated echo is that:

> "*Speaking as somebody who has been working with meta-data for years — which scientists hate with a passion — they hate it until the next time they want to use somebody else's dataset and then they whinge because there's no meta-data that describes it and tells them how to use it and what the assumptions were. Then, then they totally get it. But then it still means that they are still not happy about documenting it*" (Participant D).

This encapsulates the challenge posed at the outset of the research, namely: balancing the multiple trusts present within a context and easing tensions between the things that afford trust. Ultimately, there are no easy answers. As set out in the introductory chapter, trust brings about "all sorts of good things" (Uslaner, 2000, p.569). No one argues either in real life or in academia that trust is a bad thing, and relatively few authors speak to the necessity of 'well-placed trust'. A point made throughout this research is that too much trust, when placed in an untrustworthy trustee is something to be cautious of, but cannot specifically be avoided. It has even been argued that opportunities for failure should be designed into systems (Knowles, Rouncefield, et al., 2015). Secondly, that trust is so essential but intangible leads

people to think that trust is easy to achieve. What this research shows is that *it is not*. There aren't any shortcuts, recipes, or system features that can lead us directly to the attainment of trust. Perhaps the reason why there are so few models of trust in the socio-technical system design space is because of this. Trust cannot be reduced down to a simple model, and doing so will only lead to implementation of solutions that are doomed to fail in promoting trust.

Rather than expending effort into generating yet another trust model existing within the literature, this research has instead sought to contribute a more useful output to those working within this space. Trust affordances were created because of a frustration between the reality of trust versus the starkly linear, inappropriately rational, aseptic approaches illustrated by diagrammatic trust models. Trust affordances are ways of structuring system design prior to delving into specifications and requirements. Trust affordances developed in this dissertation can, and have, been adopted in further research, demonstrating transferability, reusability, and value to scholarly discourse. Barclay, Preece, Taylor, Radha, et al. (2023) utilise these trust affordances to identify transparency, verifiability, and accountability within machine learning (ML) "between participants in shared data and knowledge-based systems [by looking at] technologies that have been shown to provide affordances of trust between disconnected parties in different environments" (p.2). By adopting the vision of trust affordances as set out in this research, Barclay, Preece, Taylor, Radha, et al. (2023) develop an ML architecture with features that afford trust for their context, including signed certifications, a catalogue of the model from creation to present use, and a user interface to facilitate verifiability. Similarly Liao and Sundar (2022) utilise the trust affordances approach within a design scenario for artificial intelligence (AI) in a healthcare context, illustrating the need to empower users to align their placement of trust in the trustworthy and for designers to include the multiple facets of context. In adopting trust affordances derived from this research, Liao and Sundar (2022) "highlight transparency and interaction as AI systems' affordances that present a wide range of trustworthiness

cues to users" (p.1257) in order to explore the responsible communication of AI trustworthiness in order to allow for the development of appropriately well-placed trust. Both of these approaches, despite being applied in different domains, illustrate the value of trust affordances: taking a refined approach to affording trust for different stakeholders and utilising trust affordances as an element of scaffolding when formalising a system structure. Beyond this, one can also imagine other ways of applying trust affordances. For instance, reflecting on the examples in section 3.3.1 trust affordances align with all the relevant conceptualisations of affordances, promoting trust between work colleagues (social affordances) and employees and employers (organisational affordances). Again, this is not a case of adding a component of an interface and automatically generating trust. It's about reflecting multiple stakeholder trusts in dialogue at various levels throughout the design process relative to a specific context.

8.3 Future Work

This dissertation has avoided definitive statements about the trust-promoting qualities of any particular system design or component thereof. This speaks to the aim of a trust affordances approach, there is no black and white, but *shades of grey* (Thornton, Knowles, and Blair, 2021). Appreciation of nuance is the strength of this analysis. Producing a guideline or checklist holds a certain appeal, but would ultimately lead to disappointment. The central argument of this work is that such approaches do not and cannot work. This research is a call to action for system designers to forgo a narrow approach: extrapolating instead to a whole systems approach (Easterbrook, 2014; Meadows, 2008). In lieu of a magic formulae when designing for trust, the reality is more of a witch's brew wherein designers "become alchemists of trust to enable a new way of thinking and of designing" (Thornton, Knowles, and Blair, 2022, p.1395). A common theme is that there is no universal answer. What affords trust for one user might not for another, or might even cause

them to distrust. And, moreover, what works at some point in time might not work later on. There is a need to revise and adapt continuously, engaging with the social dynamics of these systems as they unfold over time. There is always a risk that final system does not fully attend to any users in any meaningful way (section 2.7), but this does not mean that we should not try. There are two avenues for future work that build on the important nuances laid out in this work. The first concerns translating the outputs of this research, trust affordances, into system design, including the specification of concrete features. This work is currently underway, exploring the design of a Virtual Lab for trustworthy and accountable decision-making. Addressing every trust issue for every stakeholder is unachievable and would restrict the opportunities to trust for others. But this remains a challenging quandary on *how* to action this research. Currently, a research design has been developed to evaluate the ways in which we can afford trust, undertaking an iterative approach using system prototypes and stakeholder engagement (Denning et al., 2010; McGonigle et al., 2020).

The second avenue for future work concerns structural changes at a community level, affecting how trust is able to unfold. For the environmental data science community and the Research Council, adoption of tools such as Virtual Labs will promote a real culture change in how environmental data science is done, thus requiring re-examination of how to buttress trust as these relationships evolve. Data curators as recipients of institutional trust from trustors will, in order to achieve trust from multiple stakeholders, have to adopt the role of trustee purposefully. This will involve taking a leading role in long-term research to transform culture, as eluded to within their strategic priorities (Natural Environment Research Council, 2022). However, the introduction of technology alone will not create this transformation. A demonstrable effort to create a culture of empathy, care, and benevolence is needed and these cues signalled to multiple stakeholders. As a subjective construct, one can easily see what trust means to oneself, but it is challenging to understand what trust means to another. A strategy to align the multiple trusts of environmental

data science stakeholders is needed, such that one can see the ways in which others place trust. Trust can be undervalued, and when proposed features were suggested to participants (chapter 6) some were reluctant to agree to these based upon time and effort required on their behalf. Working towards fulfilling trust as a trustee is in competition with taking care of more immediate needs with more immediate returns on investment, and this is compounded when the trustor is unknown. Unlike organisational trust, trust in environmental data science is unbounded both in time and space. As trustees, the trust placed in them is often invisible and intangible, making their ability to fulfil trustor's expectations challenging. The role of Research Council's as curators is not only to archive and curate data, but to take a central role in the formation and ongoing sustenance of all three dimensions of trust. Technology, when designed appropriately, can help to afford well-placed trust, but ultimately socio-technical problems require socio-technical solutions.

Appendix A

Selection of a Constructivist Paradigm

	Definition	Constructivist Position
Epistemology	What counts as knowledge; the kinds of knowledge that are useful and possible; and how knowledge claims are justified (Bryman, 2008, Creswell, 2013, Crotty, 1998, Moses and Knutsen, 2019).	"Subjectivist": Evidence based on participants' views and interpretations; context highly important; interactive and personal data collection (Charmaz, 2009, Guba and Lincoln, 1994, Mertens, 2014, Rubin and Rubin, 2005).
Ontology	The nature of the world under investigation; whether social entities can and should be considered objective entities; and whether there is a singular reality or multiple realities (Bryman, 2008, Leeming, 2018, Moses and Knutsen, 2019).	"Relativist": Belief that knowledge is socially constructed; that there is no singular 'truth' but one that differs individually, leading to multiple realities (Guba and Lincoln, 1994, Mertens, 2014, Rubin and Rubin, 2005).
Axiology	The values that researchers bring to their study, whether values are excluded and denied influence, versus the inclusion of values which are seen to have impact upon the research (Creswell, 2013, Guba and Lincoln, 1994).	Belief that values should be included and are recognised as having influence; thus consider the act of positioning oneself within the context of research as highly important (Angen, 2000, Charmaz, 2009).

Table A.1: The epistemological, ontological, and axiological positions of a constructivist paradigm.

	Social-Psychological Approach	**Constructivist Perspective**	**Case Study Methodology**	**Qualitative Methods**
Focus of study	Perceptions of the trustee's attributes and contextual factors (Blomqvist, 1997, Misztal, 2020).	Multiple perceptions of a phenomenon including the influence of context (Creswell and Miller, 2000, Guba and Lincoln, 1994).	The 'how' and 'why' of a complex phenomena in a context through a variety of lenses (Denscombe, 1998, Yin, 2017).	The 'how' and 'why' of complex concepts to develop a detailed understanding (Agee, 2009, Leeming, 2018, Schutt, 2012).
Compatibility with other approaches	Qualitative approaches preferred, avoiding imposition of a singular definition of trust (Lyon et al., 2015)	Consistent with a qualitative methodology to develop understanding (Crotty, 1998, Leeming, 2018, Mertens, 2014).	Compatible with different philosophical approaches and methods (Anfara, 2008, Bryman, 2008, Stake, 1995).	Not limited to a specific paradigm or methodology (Angen, 2000, Bryman, 2008, Maxwell, 2009).
Methods	No single method preferred, advocate for multiple methods to understand the multifaceted mature of trust (Lyon et al., 2015).	Favours interactive and personal style of data collection, supported by multiple methods, with a variety of participants (Charmaz, 2009, Creswell, 2013)	Encouraged to use a variety of data sources, data types, and data analysis methods (Boblin et al., 2013, Denscombe, 1998, Petty et al., 2012).	Using naturalistic forms of inquiry to understand participants and their context (Charmaz, 2009, Richards, 2005).

Table A.2: Connections between the approach to trust, philosophical paradigm, methodology, and methods. Columns read from left to right showing compatibly of approach and paradigm, then choice of methodology, followed by methods.

Appendix B

Participant Information Sheet, Invite, Consent Form, and Interview Schedule for Data Curator Interviews

Lancaster University

E-mail to recruit participants

Dear [name],

I am a Research Associate on the EnvChain project at Lancaster University. This project explores the potential use of blockchain-based approaches to building trust in the creation, audit and use of data. Blockchain is a developing technology that can be used to trade goods and services in a decentralised and distributed manner. Blockchain-based approaches seek to foster trust through identification and verification of producers and consumers. The use of this approach with environmental data is new, and your participation would contribute to innovation in this area. However, prior knowledge on this technology is not expected.

As part of this project, we wish to interview a range of stakeholders to identify issues of trust and auditability in environmental data transactions. This may be as a data producer, data consumer, or both. The interview will be semi-structured and will not last for more than one hour. The questions will focus on the nature of the environmental data in your role and any problems you face; if trust is relevant to the creation and use of data; and any issues that have arisen as a result of a lack of trust or audibility.

I would greatly appreciate taking the time out of your schedule to participate in my research. If you are interested in taking part in my study, or need any more information regarding my project, please do not hesitate to contact me.

Regards,

Lauren Thornton

Figure B.1: Interview invite.

Lancaster University

School of Computing and Communications

CONSENT FORM

Project Title: Envchain

Name of Researchers: Lauren Thornton Email: l.thornton2@lancaster.ac.uk

Please tick box

☐ **1.** I confirm that I have read and understand the information sheet for the study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

☐ **2.** I understand that my participation is voluntary and that I am free to withdraw at any time, without giving a reason. If I withdraw within six weeks of the commencement of the study my data will be removed. If I am involved in focus groups and then withdraw my data will remain part of the study.

☐ **3.** I understand that any information given by me may be used in future reports or presentations by the researchers, but my personal information will not be included, and I will not be identifiable.

☐ **4.** I understand that my name, my organisations name, or any personally identifiable data will not appear in any reports or presentations.

☐ **5.** I agree to the interview being audio recorded, and I understand that the interview will be transcribed, and that data will be protected on encrypted devices and kept secure.

☐ **6.** I agree to take part in the above study.

_____	_____	_____
Name of Participant	Date	Signature

I confirm that the participant was given an opportunity to ask questions about the study, and all the questions asked by the participant have been answered correctly and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and the consent has been given freely and voluntarily.

Signature of Researcher /person taking the consent _____

Date _____ (DD/MM/YYYY)

One copy of this form will be given to the participant and the original kept in the files of the researcher at Lancaster University

Figure B.2: Consent form.

Lancaster University

School of Computing and Communications

Participant information sheet

I am a Research Associate on the EnvChain project at Lancaster University and I would like to invite you to take part in a research study about the issues of trust in environmental data, and the application of blockchain technology to this domain.

Please take the time to read the following information before you decide whether or not you wish to take part.

What is the study about?

This project explores the potential use of blockchain-based approaches to building trust in the creation, audit and use of data. Blockchain is a developing technology that can be used to trade goods and services in a decentralised and distributed manner. Blockchain-based approaches seek to foster trust through identification and verification of producers and consumers. The use of this approach with environmental data is new, and your participation would contribute to innovation in this area. However, prior knowledge on this technology is not expected.

Why have I been invited?

I have approached you because I am interested in the views of academics and researchers in this domain. I believe your knowledge in the area and your views will be valuable to my study.

What will I be asked to do if I take part?

If you decided to take part, this would involve the participating in a semi-structured interview lasting roughly one hour.

What are the possible benefits from taking part?

If you could take part in this study, your insights would contribute to the understanding of trust and auditability in environmental data. This would be beneficial to the research community in terms of knowledge gained.

Do I have to take part?

No. It's completely up to you to decide whether or not you take part. Your participation is voluntary, and you are free to withdraw at any time before, during or within six weeks of the interview, without giving any reason.

What if I change my mind?

You are free to withdraw at any time and if you want to withdraw, I will extract any data you contributed to the study and destroy it. However, it is difficult and often impossible to take out data from one specific participant when this has already been anonymised or pooled together with other people's data. Therefore, you can only withdraw up to 6 weeks after taking part in the study.

Figure B.3: Participant information sheet, pages 1-3.

Lancaster University

School of Computing and Communications

What are the possible disadvantages and risks of taking part?

There are no foreseen risk or disadvantages to taking part, apart from volunteering your valuable time in participating in the interview.

Will my data be identifiable?

After the interview, only I, the researcher conducting this study, and my line managers will have access to the data you share with me. I will keep all personally identifiable information about you confidential, that is I will not share it with others. I will anonymise any audio recordings and hard copies of any data. This means that I remove any personal information.

How will my data be stored?

Your data will be stored in encrypted files (that is no-one other than me and my line managers will be able to access them) and on password-protected computers.
I will keep data that can identify you separately from non-personal information (e.g. your views on a specific topic).

How will we use the information you have shared with us and what will happen to the results of the research study?

I will use the data you have shared for academic purposes only. This will include at the presentation to NERC on the 23rd March, 2018. When writing up the findings from this study for the presentation, I would like to reproduce some of the views and ideas you shared with me. When doing so, I will only use anonymised quotes (e.g. from our interview with you), so that although I will use your exact words, you cannot be identified in our publications.

Who has reviewed the project?

This study has been reviewed and approved by the Faculty of Science and Technology Research Ethics Committee.

What if I have a question or concern?

If you have any queries or if you are unhappy with anything that happens concerning your participation in the study, please contact myself:

Lauren Thornton, l.thornton2@lancaster.ac.uk**, D22** Computing and Communications, InfoLab21, Lancaster University, Lancaster, LA1 4WA

Or one of my line managers:

Nigel Davies, Department of Computing and Communications, Lancaster University Lancaster, LA1 4WA, T: +44 (0)1524 510327.

E-mail: nigel@comp.lancs.ac.uk

Gordon Blair, C15, Computing and Communications, InfoLab21, Lancaster University, Lancaster, LA1 4WA
E-mail: g.blair@lancaster.ac.uk

And if you have any concerns or complaints that you wish to discuss with a person who is not directly involved in the research, you can also contact:

Adrian Friday, Head of Department, School of Computing and Communications, InfoLab21, Lancaster, LA1 4WA, T: +44 (0)1524 510326. E-mail: a.friday@lancaster.ac.uk

Thank you for considering your participation in this project.

Lancaster University

Interview Schedule

Please could you give me a brief overview of role and if the terms of 'data producer', 'data consumer', or 'data custodian' – or a combination of these three are the most applicable?

If a data curator, could you please describe what this entails from your perspective?

Could you tell me a little bit more about the types of environmental data that you use?

What general issues or requirements do you think exist now that you face with regards to environmental data in particular?

In general, are there any issues that spring to mind immediately when you think about trust with regards to the environmental data that you work with?
[Follow up questions tailored to answers above].

Are there any practices that you currently use to structure and convey data to a specific audience?

Are there any policies or processes at your data centre that are relevant to the role of a data curator?

Are there any issues that are pertinent to a data centre as a mechanism for trust?

Have you felt or had any issues with trust and the variety of sources? Anything that had led to a lack of trust?

Do you find there are any issues or challenges with the variety of data sources that you use or have access to?

Have you come across any issues and challenges with different data types and bringing these together?

Have you faced any issues or challenges regarding the flow of data (to or from you) in terms of trust?

Any further issues or challenges around the reliability or accuracy of data and its sources?

And if so, how you're communicating that?

Are there any further issues or challenges that we haven't discussed that you think are pertinent to trust and environmental data?

Looking to the future any tackling the issues that you have raised – do you have any requirements or needs that are based on your role or your data centres role in the future?

And, considering the systems in place currently, how do you think these will work together?

What keeps you at night – if you could change something or anything what would you want? It can be realistic or completely out of the box.

Is there anything you want to add that we haven't covered?

Finally, do you think there's potential for blockchain to work or help as part of a solution?

Figure B.4: Interview schedule.

Appendix C

Participant Information Sheet, Invite, Consent Form, and Interview Schedule for Data Producer and Data Consumer Interviews

E-mail to recruit participants

Dear [name],

I am a PhD student at Lancaster University. My project seeks to understand the role of trust when it comes to conducting environmental data science. I am interested in how trust is perceived regarding forms of secondary data, given that the literature points towards a lack of inclusion of this data within scientific research. To do so, I am looking for researchers, academics and scientists who have experience in this field and would be willing to provide some insight into this area. I hope to develop requirements necessary for trusted data systems, and in particular to contribute to the design of virtual labs, which are collaborative research environments.

As a person involved in this field, I am writing to enquire about participating in my project. This research will take the form of a semi-structured interview of roughly one hour. I have completed the ethical procedures required by the University, and your participation would remain anonymous and unidentifiable in both my thesis and any publications I produce.

I would greatly appreciate taking the time out of your schedule to participate in my research. If you are interested in taking part in my study, or need any more information regarding my project, please do not hesitate to contact me.

Regards,
Lauren Thornton

Figure C.1: Interview invite.

Lancaster University

CONSENT FORM

Project Title: Data, Decisions, and Doubt: Exploring Trust in Environmental Data Science

Name of Researchers: Lauren Thornton Email: l.thornton2@lancaster.ac.uk

Please tick each box:

☐ **1.** I confirm that I have read and understand the information sheet for the study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

☐ **2.** I understand that my participation is voluntary and that I am free to withdraw at any time, without giving a reason. If I withdraw within six weeks of the commencement of the study my data will be removed. If I am involved in focus groups and then withdraw my data will remain part of the study.

☐ **3.** I understand that any information given by me may be used in future reports, articles or presentations by the researchers, but my personal information will not be included, and I will not be identifiable.

☐ **4.** I understand that my name, my organisations name, or any personally identifiable data will not appear in any reports, articles or presentations.

☐ **5.** I understand that data regarding my organisational affiliation and role will be retained, but this information will be kept anonymous in any publications resulting from this research unless I have given express consent otherwise.

☐ **6.** I agree to the interview being audio recorded, and I understand that the interview will be transcribed, and that data will be protected on encrypted devices and kept secure.

☐ **7.** I understand that data will be kept according to University guidelines for a minimum of 10 years after the end of the study.

☐ **8.** I understand that all research data stored at Lancaster University is covered under the EU General Data Protection Regulation (GDPR) and the UK Data Protection Act (2018).

☐ **9.** I agree to take part in the above study.

_____ _____ _____

Name of Participant Date Signature

I confirm that the participant was given an opportunity to ask questions about the study, and all the questions asked by the participant have been answered correctly and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and the consent has been given freely and voluntarily.

Signature of Researcher /person taking the consent _____

Date _____ (DD/MM/YYYY)

One copy of this form will be given to the participant and the original kept in the files of the researcher at Lancaster University

Figure C.2: Consent form.

Lancaster University

Participant information sheet

For further information about how Lancaster University processes personal data for research purposes and your data rights please visit our webpage: www.lancaster.ac.uk/research/data-protection

I am a PhD student at Lancaster University, and I would like to invite you to take part in a research study about the trustworthiness of secondary data and data science techniques in environmental science. I am interested in studying mechanisms to foster trust and how these may be utilised in creating enhanced virtual labs.

Please take the time to read the following information before you decide whether or not you wish to take part.

What is the study about?
This study aims to explore trust in data. I am investigating how lack of trust potentially results in the limited use of this data in scientific research and how trust is fostered. Additionally, I am interested in designing solutions to overcome this lack of trust. This includes how to communicate certainties and uncertainties present in the data; what the most effective presentation for this communication would be; and if implemented, whether these solutions would indeed foster trust and use of these data sources. I am hoping to contribute to designing enhanced virtual labs, which are collaborative research environments and may be beneficial to the production of knowledge in environmental sciences.

Why have I been invited?
I have approached you because I am interested in the views of academics and researchers in this domain. I believe your knowledge in the area and your views will be valuable to my study.

What will I be asked to do if I take part?
If you decide to take part, this would involve participating in a semi-structured conversational interview lasting roughly one hour. Every effort will be made for interviews to take place in a location that is convenient for you - either on Lancaster University campus, off campus, or via Skype – and will be agreed upon prior to participation.

For some participants, you may also be asked to be involved in a small activity drawing the components of virtual labs. You will be asked beforehand if you are comfortable participating in the activity. The activity will involve using pieces of paper and other materials to construct the elements that you feel should be included in a collaborative research environment. If undertaken, this activity will be conducted within the interview, and therefore participation should for roughly one hour and should not last more than one and a half hours. However, if this is not possible e.g. it is more beneficial to you, I will conduct the activity separate to the interview. Therefore, totalling participation of no more than two hours. There may also be a follow up to the activity, presenting the culmination of my findings back to participants – this will be sent to you via email, and again it is your decision to respond to this and to communicate back via email with your thoughts and opinions on my results.

What are the possible benefits from taking part?
If you could take part in this study, your insights would contribute to the understanding of trust in computing and in secondary sources of data. This would be beneficial to the research community in terms of knowledge gained; and in the future

Figure C.3: Participant information sheet, pages 1-3.

Lancaster University

the potential inclusion of more data in research projects. In the event this research leads to the development of data system prototypes, you may be a recipient of them for use in your work.

Do I have to take part?
No. It's completely up to you to decide whether or not you take part. Your participation is voluntary, and you are free to withdraw at any time before, during or within six weeks of the interview, without giving any reason.

What if I change my mind?
You are free to withdraw at any time and if you want to withdraw, I will extract any data you contributed to the study and destroy it. However, it is difficult and often impossible to take out data from one specific participant when this has already been anonymised or pooled together with other people's data. Therefore, you can only withdraw up to 6 weeks after taking part in the study.

What are the possible disadvantages and risks of taking part?
There are no foreseen risk or disadvantages to taking part, apart from volunteering your valuable time in participating in the interview.

Will my data be identifiable?
After the interview, only I, the researcher conducting this study, and my dissertation supervisors will have access to the data you share with me. I will keep all personally identifiable information about you confidential, that is I will not share it with others. I will anonymise any audio recordings and hard copies of any data. This means that I remove any personal information.

How will my data be stored?
Your data will be stored in encrypted files (that is no-one other than me and my supervisors will be able to access them) and on password-protected computers.
I will keep data that can identify you separately from non-personal information (e.g. your views on a specific topic).

How will we use the information you have shared with us and what will happen to the results of the research study?
I will use the data you have shared for academic purposes only. This will include my PhD thesis and other academic publications, such as journal articles. I may also present the results of my study at academic conferences.

When writing up the findings from this study, I would like to reproduce some of the views and ideas you shared with me. When doing so, I will only use anonymised quotes (e.g. from my interview with you), so that although I will use your exact words, you cannot be identified in publications.

Who has reviewed the project?
This study has been reviewed and approved by the Faculty of Science and Technology Research Ethics Committee at Lancaster University.

What if I have a question or concern?

If you have any queries or if you are unhappy with anything that happens concerning your participation in the study, please contact myself:

Lauren Thornton, l.thornton2@lancaster.ac.uk

C22, Computing and Communications, InfoLab21, Lancaster University, Lancaster, LA1 4WA

Or one of my supervisors:

Bran Knowles, C45, Computing and Communications, InfoLab21, Lancaster University, Lancaster, LA1 4WA
E-mail: b.h.knowles1@lancaster.ac.uk

Gordon Blair, C15, Computing and Communications, InfoLab21, Lancaster University, Lancaster, LA1 4WA
E-mail: g.blair@lancaster.ac.uk

If you have any concerns or complaints that you wish to discuss with the institute funding this work, you can also contact either of the Co-Directors of the Data Science Institute:

Professor Nigel Davies, Department of Computing and Communications, Lancaster University Lancaster, LA1 4WA, T: +44 (0)1524 510327.
E-mail: nigel@comp.lancs.ac.uk

Professor Idris Eckley, Department of Mathematics & Statistics, Lancaster University, LA1 4WA, T: +44 (0)1524 593066.
E-mail: i.eckley@lancaster.ac.uk

And if you have any concerns or complaints that you wish to discuss with a person who is not directly involved in the research, you can also contact:

Adrian Friday, Head of Department, School of Computing and Communications, InfoLab21, Lancaster, LA1 4WA, T: +44 (0)1524 510326. E-mail: a.friday@lancaster.ac.uk

Thank you for considering your participation in this project.

Draft Interview Schedule

Questions to be asked once participant information sheet has been given and/or I have introduced myself and asked if they would like to participate.

Draft Interview Questions:
Could you tell me about the environmental data you use? What types of environmental science?
Do you predominantly use primary or secondary sources of data, or a combination of both?
What are the predominant issues you face with the data you use?
Do you use data for research or are you more involved with archival/curation?
Have you ever thought about trust in relation to your use of data?
Have you ever considered the impacts that data can have on results?
What springs to mind when you think about: trusting data, and trustworthy data?
What springs to mind when you think about: not trusting data, and untrustworthy data?
Thinking about having trust in data: in what ways do you come to this position? E.g. is there anything in particular you look for?
Have you ever had the experience of getting or collecting data and not trusting it?
Picking out different elements of trust in data - do you have any thoughts on:

- Quality?
- Accuracy?
- Traceability?
- Discoverability?
- Uncertainty?

Are there any current mechanisms you utilise to ensure trustworthy data? E.g. get it from particular places or look for particular things within it?
What would you need or like to see as mechanisms to trust data? E.g. Meta-data/Additional information?

Introduce concept of collaborative research environments if applicable to case.
In your line of work/research do you feel any need for such an environment?
What key characteristics do you feel it would need to have?
Do you feel this would benefit your work with environmental data? In what ways?
Do you think this would help with the predominant issues you outlined with data?
Do you feel that this would be beneficial to having trust in data and trustworthy data?
And finally, is there anything else that you feel we have not discussed?

Thank you.

Figure C.4: Interview schedule.

Appendix D

Consent Form and Survey Schedule for Data Producer and Data Consumer Survey

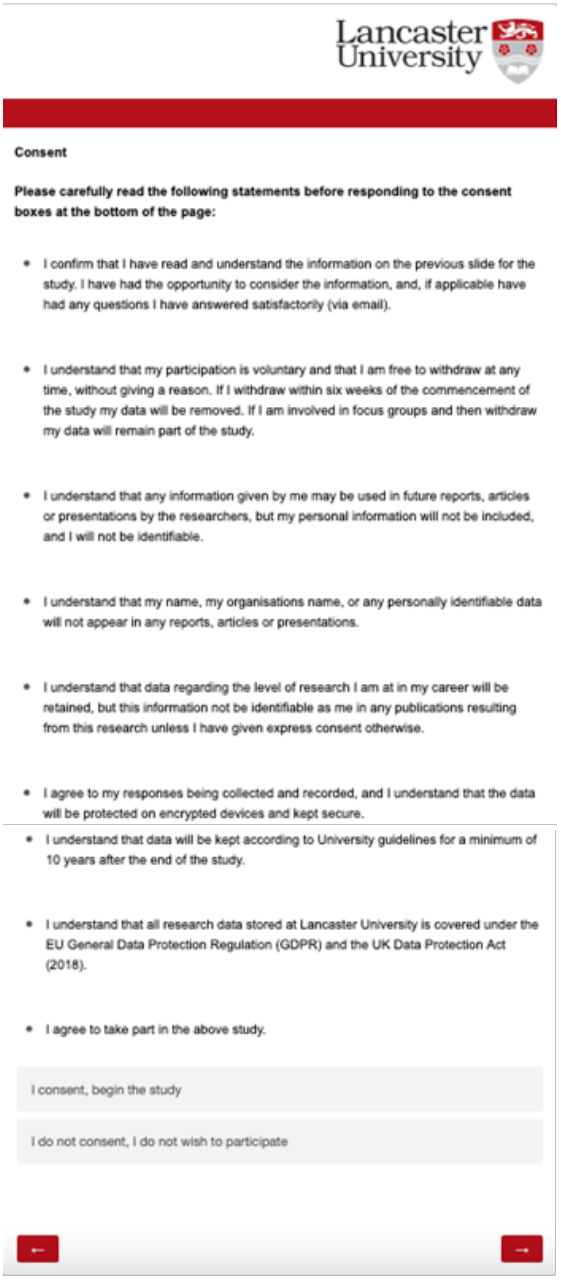


Figure D.1: Consent form.

Questionnaire contents

Demographics:
I am currently a: [please select the most appropriate]

- Doctoral Research Student
- Post-Doctoral/Early Career Researcher
- Senior Researcher
- Lecturer
- Manager
- Non-academic role, e.g. Technical and Development
- Other [blank text option]
- Prefer not to say

The questions asked in the questionnaire are:
First Section: Statements - General
Regarding the following statements how do you feel? [Scale: Strongly Agree – Agree – Neither Agree or Disagree – Disagree – Strongly Disagree]
Trust and trustworthiness **is important** for environmental data science as a whole. [Scale]
Trust and trustworthiness **is important** for me in my research/role. [Scale]
I have **never considered** trust or had a reason to consider trust during my research/role. [Scale]
Trustworthy by design is an **important goal** when designing and implementing technology within environmental data science. [Scale]
When conducting my research, for instance using others' data or models, I **have considered** the trustworthiness of the object or those who have created it before I have used it. [Scale]
Feeling informed and having an understanding is important to me when placing trust. [Scale]
Please choose which statement best fits your opinion within your research: [Choice]

- Trust to me is primarily associated with people and relationships.
- Trust to me is primarily associated with data, models and research outputs.
- Both, trust to me is associated with a mixture of social and technical.
- Trust to me is associated with none of these.

In your own words, how would you define something as trustworthy within environmental data science? [Free form]
In your own words, how would you describe what trust means to you when thinking about environmental data science? [Free form]
How would you seek to determine the trustworthiness of another person or of data and/or models? E.g. reputation, documentation. And what are the things you look for/frequently find missing?

Second Section: Trust Affordances and Virtual Labs
Thinking about the features and properties of technology, in this instance thinking about virtual labs, and **how these relate to enhancing trust**, please rate these statements:
Provenance information is useful to me in my research/role. [Scale]
Knowing the identity of others is important to me in my research/role. [Scale]

Figure D.2: Survey schedule, pages 1-2.

Knowing the provenance of a dataset or a model is useful to me in determining trustworthiness. [Scale]
I have experience of using others' meta-data and supplementary within my research [Scale]
I have a positive experience of using the above and have no issues in using it [Scale]
Documentation always gives me all of the information I need [Scale]
I have experience of having to contact someone for additional information that was not included initially [Scale]
I often find I need additional information but have not been able to identify who to contact or have received a response [Scale]
Documentation is important to me in assessing the trustworthiness of something before I have used it [Scale]
Transparency is important in environmental data science [Scale]
I believe that features such as modified access through credentials are important to research [Scale]
How would you define transparency within environmental data science and how do you think this should be controlled/implemented? [Free form]
Having trust in others is important for me when collaborating (initially and continually) [Scale]
What factors are important to you when trusting others within a collaboration that you **already have a relationship with**? (Please put 'none' if you **do not** think any factors are important to you) [Free form]
What factors are important to you when trusting others within a collaboration that you **do not already have a relationship with**? (Please put 'none' if you **do not** think any factors are important to you). [Free form]
Are there any features of properties of technology, particularly Virtual Labs, that would help you to place trust or assess trustworthiness when working with others? (Either your own team, another discipline, or those outside of academia) [Free form]
Imagining that you are collaborating with others' from a different scientific discipline, what are the things that you would need to trust, e.g. feel informed and enable understanding? (Please put 'none' if you **do not** think any factors are important to you). [Free form]
How do you feel about Virtual Labs and their ability to promote well-placed trust? E.g. do you agree or disagree that they may help with current trust issues? [Free form]

Final Section
Is there anything else you would like to add?
If you would like to be invited to take part in future research please submit your email address here. **You are under no obligation to do so, and your email address will not be linked to your answers within this questionnaire**.

Appendix E

Sticky Notes and Group Feedback Generated Through the Data Curator Workshop

Session One exercise feedback

Things that have really resonated with you:

- Standardisation – common vocabularies needed
- Provenance needed in multi-discipline collaborations
- Need very clear provenance networks for sensor data
- Don't get quality data and contextual meta-data even when it's funded by NERC
- Getting provenance information is a lot of work, in terms of modelling etc. too much work in getting that information, unless its automatic or employ specific job roles
- Want the persistence of some information but not others
- Data centre can only deal with the data and associated information they get, not what happens before
- When you formalise provenance you potentially add a lot of complexity – which is passed onto the user – don't want to pass that on to business users for example

Things you haven't heard:

- Complexity, challenges of adding complexity to the user – how blockchain implementation needs to be invisible/not in the user's face
- Volumes of data – in terms of adding extra, i.e. provenance
- Cost savings – implementation of blockchain, any sense of major senses and major costs
- Where does provenance start – where does the data centre stand on that chain?
- Blockchain needs to be user driven, but defining 'user' is not straightforward
- Culture change within science – biggest problem, long time to change
- Start at consumer end – it will feed into custodian and provider anyway – it is a chain
- Intellectual property, data licencing
- Gulf between IT and science communities in terms of how they handle the data
- Impact of having more information on data – not fully aware of issues even with information, if not from community – if problems have been highlighted, will it lead to less reuse?
- Encourage data producers' behaviour to be active in uploading and information
- Obama law – data has to be traceable if in policy, need for something similar in the UK?
- "If stepping out of area (i.e., outside of a specific research community or science discipline), do I start from level of trust or do I look to blockhain? – I would probably look towards blockchain that could provide transparency and provenance"
- Blockchain makes people feeling that they commit, this makes them nervous about that, and potentially provide less meta-data
- Want to sit with academics rather than what you actually get in data sets – side-channel to get a richness of data – need for people-facing services

Figure E.1: Typed up notes of the feedback following workshop presentation of the six themes found in qualitative interviews.

Figure E.2: Sticky notes collected from the workshop sessions, arranged by role: producer, curator, consumer.

Appendix F

Preliminary Organisation of Data Producer and Data Consumer Interview Data into Themes

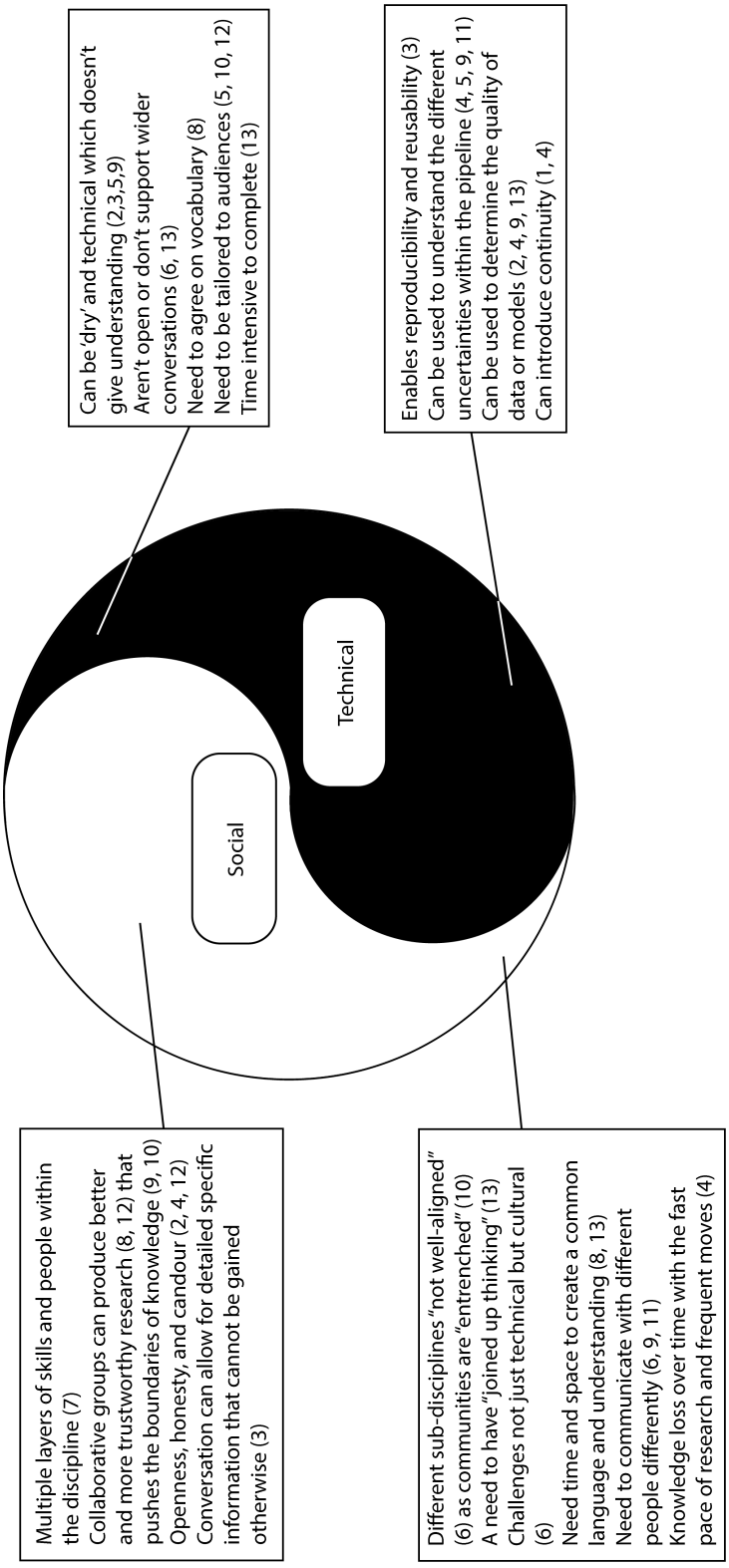
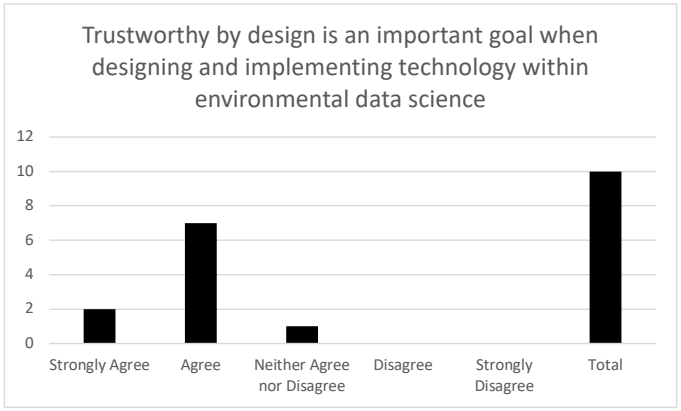


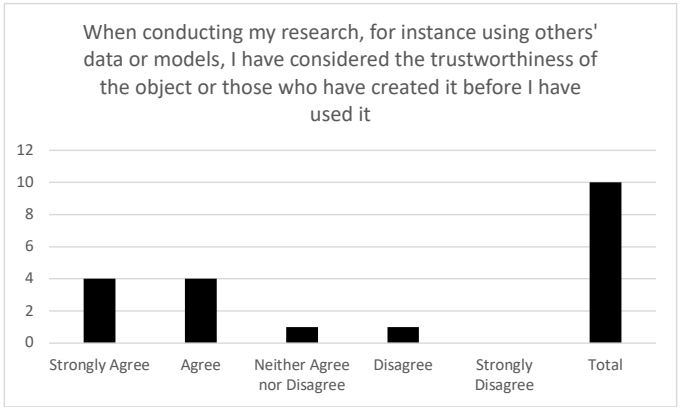
Figure F.1: A copy of preliminary organisation of social and technical themes from the interview data, drawn up using Adobe Photoshop.

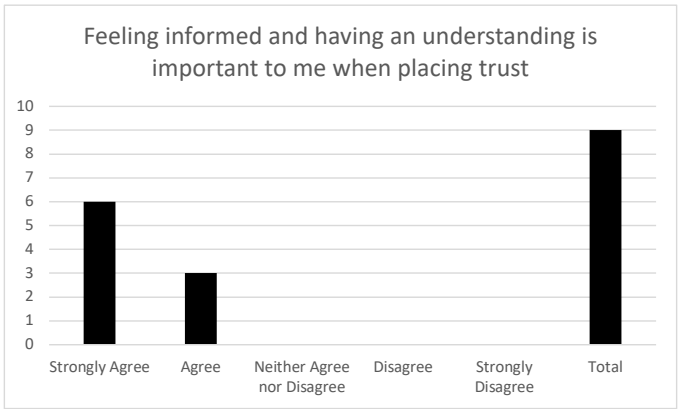
Appendix G

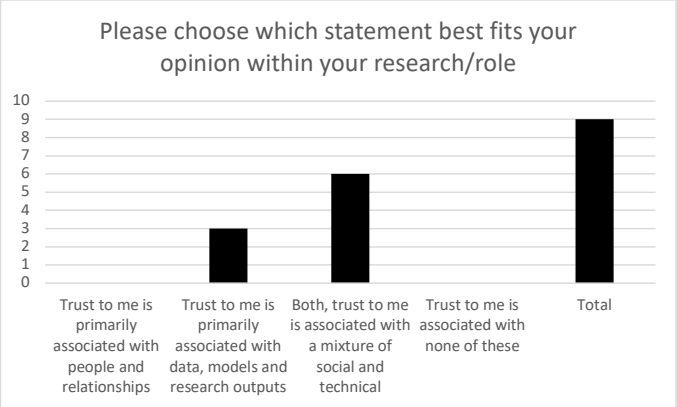
Respondent Data Generated Through the Data Producer and Data Consumer Survey

Figure G.1: Recorded answers to the Qualtrics survey for the Centre of Excellence case, pages 1-8.


Trustworthy by design is an important goal when designing and implementing technology within environmental data science
12
10
8
6
4
2
0
Strongly Agree
Agree
Neither Agree nor Disagree
Disagree
Strongly Disagree
Total


When conducting my research, for instance using others' data or models, I have considered the trustworthiness of the object or those who have created it before I have used it
12
10
8
6
4
2
0
Strongly Agree
Agree
Neither Agree nor Disagree
Disagree
Strongly Disagree
Total


Feeling informed and having an understanding is important to me when placing trust
10
9
8
7
6
5
4
3
2
1
0
Strongly Agree
Agree
Neither Agree nor Disagree
Disagree
Strongly Disagree
Total

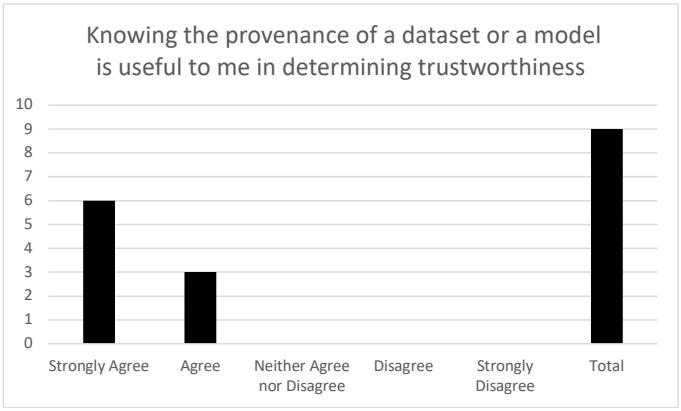


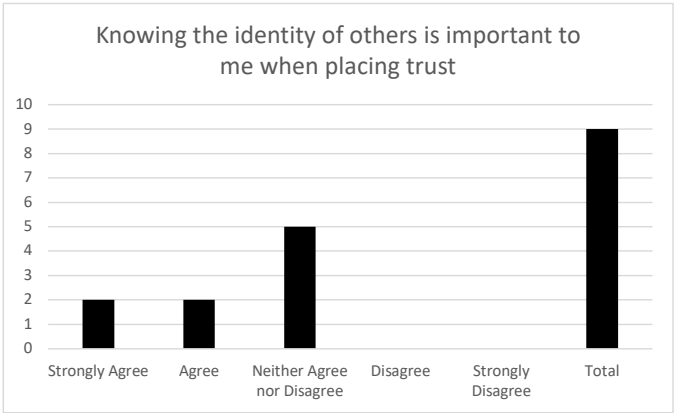
In your own words, how would you describe what trust means to you when thinking about environmental data science? How would you define something as trustworthy? [Free form]

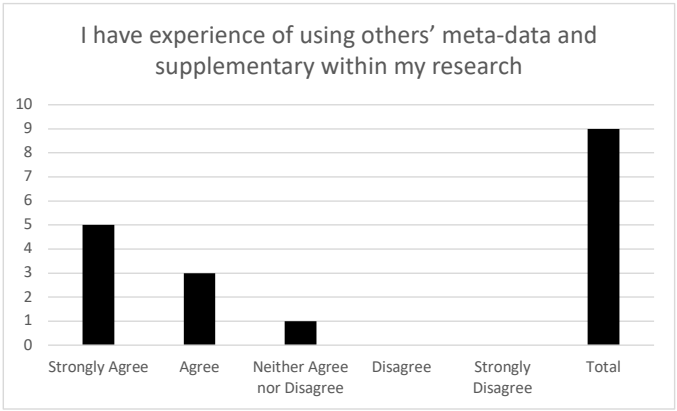
- Know where the data came from, that the research was carried out 'correctly', and all processes are transparent
- Trust is genuine, original with known sources, clear, understandable, reproducible
- For data, trustworthy means that the uncertainties are clearly documented
- Trusting that data collection has been carried out in a rigorous manner and reporting of results has been done honestly
- Availability of background information, provenance of data combined with organisation reputation for standards
- Representing science with nuance and honesty about the unknowns, knowledge gaps and uncertainties
- A model or dataset is trustworthy if I believe it has been developed using robust, peer-reviewed, validated methods, and has been demonstrated that it is reliable (or has a clear estimation of the uncertainty)

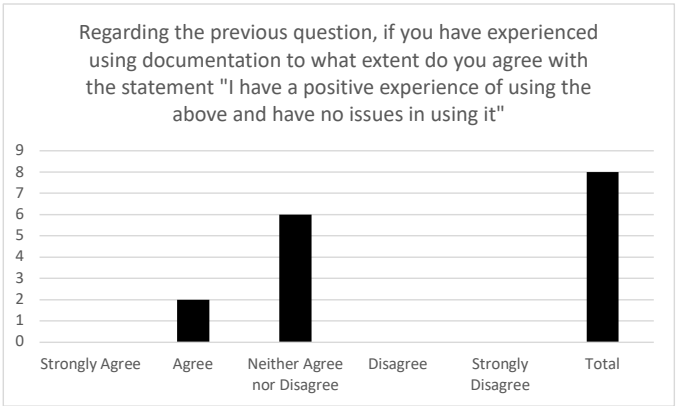
How would you seek to determine the trustworthiness of another person or of data and/or models? E.g. reputation, documentation. And what are the things you look for/frequently find missing? [Free form]

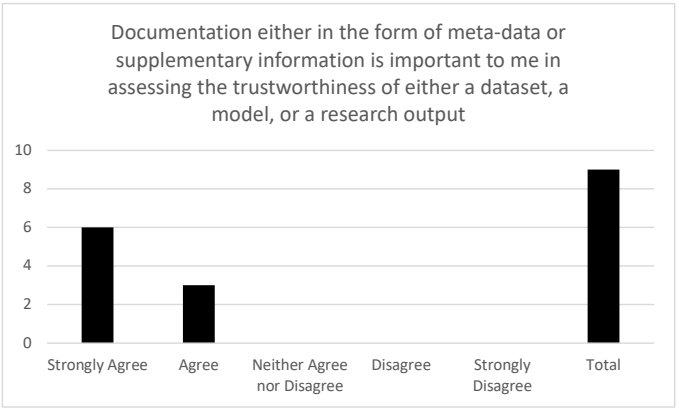
- Reproducible research. Open research practices. Willingness to engage about their research.
- Academics reputation, profile, institution and publications, technical documentation, reproducibility documentation, citations/ reuse, related publications. Documentation and reproducibility are the frequently missing
- Transparency (e.g. open access), reputation, wide use (models and data). Sometimes model documentation is missing or incomplete and I find that very frustrating.
- Read their work to determine for myself. Ask a more senior colleague who may have more experience of this person and/or models
- Ease of obtaining the data/models together with complete, well-constructed documentation
- With models- checking the code ideally! documentation is often behind. Having citations to support the science behind models and accompanying notes on limitations/confidence
- Publication (only peer-reviewed is trustworthy), track-record of group who has developed the model/dataset, reputation of lead scientist and/or institution, clear user guide documentation. The latter is often missing or difficult to find.

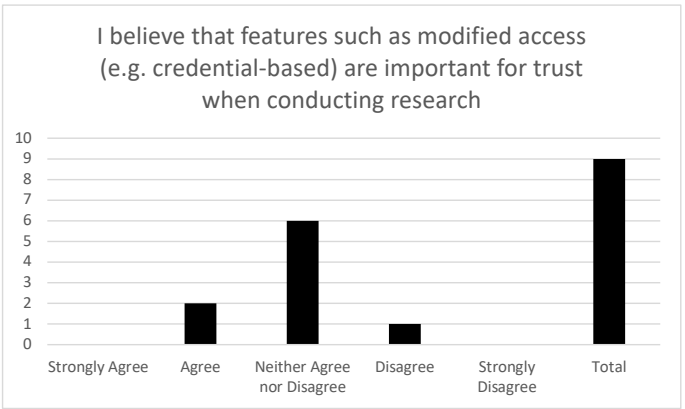

Knowing the provenance of a dataset or a model is useful to me in determining trustworthiness
10
9
8
7
6
5
4
3
2
1
0
Strongly Agree
Agree
Neither Agree nor Disagree
Disagree
Strongly Disagree
Total

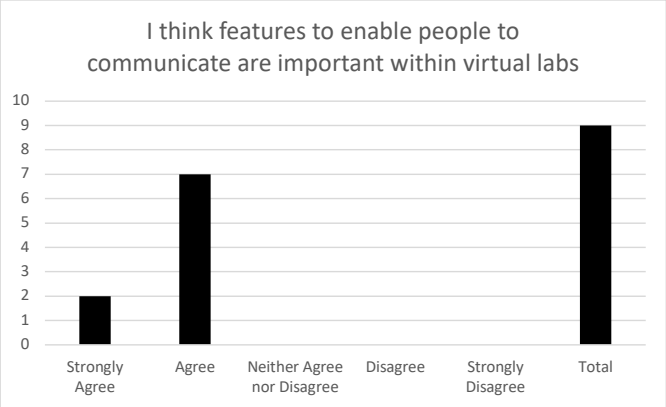

Knowing the identity of others is important to me when placing trust
10
9
8
7
6
5
4
3
2
1
0
Strongly Agree
Agree
Neither Agree nor Disagree
Disagree
Strongly Disagree
Total


I have experience of using others' meta-data and supplementary within my research
10
9
8
7
6
5
4
3
2
1
0
Strongly Agree
Agree
Neither Agree nor Disagree
Disagree
Strongly Disagree
Total


Regarding the previous question, if you have experienced using documentation to what extent do you agree with the statement "I have a positive experience of using the above and have no issues in using it"
9
8
7
6
5
4
3
2
1
0
Strongly Agree
Agree
Neither Agree nor Disagree
Disagree
Strongly Disagree
Total

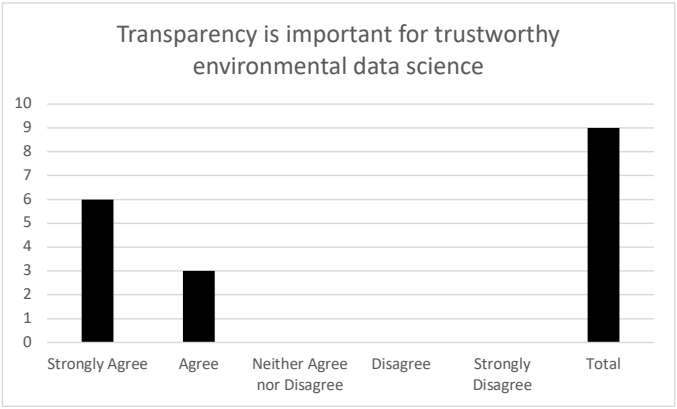

Documentation either in the form of meta-data or supplementary information is important to me in assessing the trustworthiness of either a dataset, a model, or a research output
10
8
6
4
2
0
Strongly Agree
Agree
Neither Agree nor Disagree
Disagree
Strongly Disagree
Total


I believe that features such as modified access (e.g. credential-based) are important for trust when conducting research
10
9
8
7
6
5
4
3
2
1
0
Strongly Agree
Agree
Neither Agree nor Disagree
Disagree
Strongly Disagree
Total



Are there any features for communication that you think are relevant? e.g. direct messages, specific message boards, general bulletin boards [Free form]

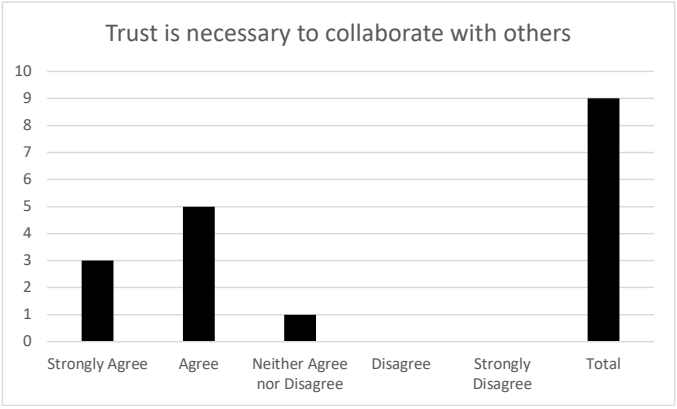
- Message boards and bulletin boards could be useful. No direct messages as virtual labs are all about transparency?
- Direct messages, specific message boards, opportunities for comments e.g. with code commits, or model/code readme
- Being able to contact the original poster is crucial



How would you define transparency within environmental data science? e.g. do you believe everything should be open and accessible to all or do you feel that there should be some form of control? [Free form]

- I think that if something is unavailable (perhaps so that some analysis can be done/published first by the data gatherer) then there needs to be an explanation of why in the documentation
- Everything should be open
- Most things should be open but there will always be some things that need access control for at least part of the lifecycle
- Openness is desirable, but clarity about all the steps in the workflow is more important
- Depends on the user. information to understand the implications of issues is more useful than the full details of the issues for some users

- I think everything should be accessible to all, but having a trace of who has used it for what purpose is a good feedback to have to demonstrate impact



What factors are important to you when trusting others within a collaboration with people you know? [Free form]

- Good communication
- Knowledge of previous work, sometimes a more personal connection
- Knowing that they are competent and are effective communicators
- Past experience
- Knowing that they are reliable (will produce deliverable in a timely manner), and their expertise

What factors are important to you when trusting others within a collaboration with people you do not know? [Free form]

- Transparency, sharing
- Knowledge of previous work
- Trusting their reputation or the opinions of them from people I do know and trust
- Knowing they are affiliated with a trustworthy organisation
- Reputation
- Their individual and institutional reputation, list of publication, track record

Are there any features of properties of technology, particularly Virtual Labs, that would help you to place trust or assess trustworthiness when working with others? (Either your own team, another discipline, or those outside of academia) [Free form]

- Transparency, sharing, access control
- I don't know
- No
- Maybe ability to score collaborators, to see whether I can trust someone based on other people's opinion? A bit like on eBay or LinkedIn or ResearchGate? (just throwing in this idea, not sure it would work in virtual labs, as it implies others being willing to "endorse" other users)

How do you feel about Virtual Labs and their ability to promote well-placed trust? e.g. do you agree or disagree that they may help with current trust issues? [Free form]

- Agree
- I think they promote transparency within data science which is good for trust
- Virtual labs make it easier to collaborate, to share data and share code. So transparency is enhanced
- The lack of transparency can be an issue

Do you think that virtual labs will be beneficial for creating a community based on trust? [Free form]

- Yes, but it will take time like any other new technology when emerging; i.e. cultural shift
- Yes
- Unsure
- If they get more widely used and accepted, they could create a community based on trust
- Cultural shift might be needed for wider adoption of virtual labs. Reluctance to change the way of working might be another barrier, on top of the issue of trust

References

Ackerman, Mark S. (2000). "The Intellectual Challenge of CSCW: The Gap between Social Requirements and Technical Feasibility". In: *Human-Computer Interaction* 15.2, pp. 179–203. DOI: 10.1207/S15327051HCI1523_5.

Ackerman, Mark S. et al. (2013). "Sharing Knowledge and Expertise: The CSCW View of Knowledge Management". In: *Computer Supported Cooperative Work* 22.4–6, pp. 531–573. DOI: 10.1007/s10606-013-9192-8.

Agee, Jane (2009). "Developing Qualitative Research Questions: A Reflective Process". In: *International Journal of Qualitative Studies in Education* 22.4, pp. 431–447. DOI: 10.1080/09518390902736512.

Anderson, Bob and Wes Sharrock (1992). "Can Organisations Afford Knowledge?" In: *Conference on Computer Supported Cooperative Work*. Vol. 1. New York, NY, USA: Association for Computing Machinery, pp. 143–161. DOI: 10.1007/BF00752436.

Anderson, Chad and Daniel Robey (2017). "Affordance Potency: Explaining the Actualization of Technology Affordances". In: *Information and Organization* 27.2, pp. 100–115. DOI: 10.1016/j.infoandorg.2017.03.002.

Anfara, Vincent A. Jr. (2008). "Theoretical Frameworks". In: *The SAGE Encyclopedia of Qualitative Research Methods*. Ed. by Lisa M. Given. Thousand Oaks, CA: SAGE Publications, pp. 869–873. ISBN: 9781412941631.

Angen, Maureen Jane (2000). "Evaluating Interpretive Inquiry: Reviewing the Validity Debate and Opening the Dialogue". In: *Qualitative Health Research* 10.3, pp. 378–395. DOI: 10.1177/104973230001000308.

Arnold, Matthew et al. (2019). "FactSheets: Increasing Trust in AI Services through Supplier's Declarations of Conformity". In: *IBM J. Res. Dev.* 63.4/5, pp. 1–13. DOI: 10.1147/JRD.2019.2942288.

Artz, Donovan and Yolanda Gil (2007). "A Survey of Trust in Computer Science and the Semantic Web". In: *Journal of Web Semantics* 5.2. Software Engineering and the Semantic Web, pp. 58–71. ISSN: 1570-8268. DOI: https://doi.org/10.1016/j.websem.2007.03.002.

Arya, Vijay et al. (2019). "One Explanation Does Not Fit All: A Toolkit and Taxonomy of AI Explainability Techniques". In: *INFORMS 2021*. Accessed 07/10/2022. DOI: https://doi.org/10.48550/arXiv.1909.03012. URL: https://research.ibm.com/publications/one-explanation-does-not-fit-all-a-toolkit-and-taxonomy-of-ai-explainability-techniques.

Ashoori, Maryam and Justin D. Weisz (2019). "In AI We Trust? Factors That Influence Trustworthiness of AI-infused Decision-Making Processes". In: arXiv: 1912.02675.

Baier, Annette (1986). "Trust and Antitrust". In: *Ethics* 96.2, pp. 231–260. DOI: 10.1086/292745.

Baker, Karen S. and Lynn Yarmey (2009). "Data Stewardship: Environmental Data Curation and a Web-of-Repositories". In: *International Journal of Digital Curation* 4 (2), pp. 12–27. URL: https://ijdc.net/ijdc/article/view/90.

Bannon, Liam (2011). "Reimagining HCI: Toward a More Human-Centered Perspective". In: *Interactions* 18.4, pp. 50–57. DOI: 10.1145/1978822.1978833.

Bannon, Liam J. and Kjeld Schmidt (1989). "CSCW: Four Characters in Search of a Context". In: *Proceedings of the First European Conference on Computer Supported Cooperative Work*. ECSCW 1989. Accessed 04/03/2021. Springer. URL: https://dl.eusset.eu/handle/20.500.12015/2602.

Barber, Bernard (1983). *The Logic and Limits of Trust*. New Brunswick, NJ: Rutgers University Press. ISBN: 9780813510026.

Barclay, Iain, Alun Preece, Ian Taylor, Swapna Krishnakumar Radha, et al. (2023). "Providing Assurance and Scrutability on Shared Data and Machine Learning Models with Berifiable Credentials". In: *Concurrency and Computation: Practice and Experience* 35.18, pp. 1–17. DOI: https://doi.org/10.1002/cpe.6997.

Barclay, Iain, Alun D. Preece, Ian J. Taylor, and Dinesh C. Verma (2019). "Towards Traceability in Data Ecosystems using a Bill of Materials Model". In: *11th International Workshop on Science Gateways*. Accessed 19/05/2021. eprint: https://ceur-ws.org/Vol-2975/paper3.pdf.

Bardram, Jakob E. and Steven Houben (2018). "Collaborative Affordances of Medical Records". In: *Computer Supported Cooperative Work* 27.1, pp. 1–36. DOI: 10.1007/s10606-017-9298-5.

Barker, Michelle et al. (2019). "The Global Impact of Science Gateways, Virtual Research Environments and Virtual Laboratories". In: *Future Generation Computer Systems* 95, pp. 240–248. DOI: https://doi.org/10.1016/j.future.2018.12.026.

Bates, Jo, Yu-Wei Lin, and Paula Goodale (2016). "Data Journeys: Capturing the Socio-material Constitution of Data Objects and Flows". In: *Big Data & Society* 3.2. DOI: 10.1177/2053951716654502.

Bauer, Henry H. (1990). "Barriers Against Interdisciplinarity: Implications for Studies of Science, Technology, and Society (STS)". In: *Science, Technology, & Human Values* 15.1, pp. 105–119. DOI: 10.1177/016224399001500110.

Baxter, Gordon and Ian Sommerville (2011). "Socio-technical Systems: From Design Methods to Systems Engineering". In: *Interacting with Computers* 23.1, pp. 4–17. DOI: 10.1016/j.intcom.2010.07.003.

Baxter, Pamela and Susan Jack (2008). "Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers". In: *The Qualitative Report* 13.4, pp. 544–559. DOI: 10.46743/2160-3715/2008.1573.

Beagrie, Neil (2006). "Digital Curation for Science, Digital Libraries, and Individuals". In: *International Journal of Digital Curation* 1, pp. 3–16. URL: https://ijdc.net/index.php/ijdc/article/view/2.

Bechhofer, Sean et al. (2013). "Why Linked Data is Not Enough for Scientists". In: *Future Generation Computer Systems* 29.2, pp. 599–611. DOI: https://doi.org/10.1016/j.future.2011.08.004.

Benbasat, Izak and Weiquan Wang (2005). "Trust In and Adoption of Online Recommendation Agents". In: *J. Assoc. Inf. Syst.* 6.3. Accessed 29/12/2021, p. 4. URL: https://aisel.aisnet.org/jais/vol6/iss3/4/.

Berners-Lee, Mike (2019). *There Is No Planet B: A Handbook for the Make or Break Years*. New York: Cambridge University Press. ISBN: 9781108337519.

Beven, Keith (2009). *Environmental Modelling: An Uncertain Future?* CRC Press. ISBN: 9780415457590.

Bierhoff, Hans-Werner and Bernd Vornefeld (2016). "The Social Psychology of Trust with Applications in the Internet". In: *Opportunities and Restrictions for Online Cooperation*. Ed. by Bernd Lahno and Uwe Matzat. Berlin, Boston: De Gruyter Oldenbourg, pp. 48–62. DOI: 10.1515/9783110508666-004.

Birnholtz, Jeremy P. and Matthew J. Bietz (2003). "Data at Work: Supporting Sharing in Science and Engineering". In: *Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work*. GROUP '03. Sanibel Island, Florida, USA: Association for Computing Machinery, pp. 339–348. ISBN: 1581136935. DOI: 10.1145/958160.958215.

Blair, Gordon, Richard Bassett, et al. (2021). "The Role of Digital Technologies in Responding to the Grand Challenges of the Natural Environment: The Windermere Accord". In: *Patterns* 2.1. DOI: 10.1016/j.patter.2020.100156.

Blair, Gordon S., Keith Beven, et al. (2019). "Models of Everywhere Revisited: A Technological Perspective". In: *Environmental Modelling & Software* 122. DOI: https://doi.org/10.1016/j.envsoft.2019.104521.

Blair, Gordon S., Amber Leeson, et al. (2019). "Data Science of the Natural Environment: A Research Roadmap". In: *Frontiers in Environmental Science* 7.121, pp. 1–14. DOI: 10.3389/fenvs.2019.00121.

Blomqvist, Kirsimarja (1997). "The Many Faces of Trust". In: *Scandinavian Journal of Management* 13.3, pp. 271–286. DOI: https://doi.org/10.1016/S0956-5221(97)84644-1.

Boblin, Sheryl L. et al. (2013). "Using Stake's Qualitative Case Study Approach to Explore Implementation of Evidence-Based Practice". In: *Qualitative Health Research* 23.9, pp. 1267–1275. DOI: 10.1177/1049732313502128.

Boersema, Jan J. (2009). "Environmental Sciences, Sustainability, and Quality". In: *Principles of Environmental Sciences*. Ed. by Jan J. Boersema and Lucas Reijnders. Dordrecht: Springer Netherlands, pp. 3–14. ISBN: 978-1-4020-9158-2. DOI: 10.1007/978-1-4020-9158-2_1.

Borgman, Christine L. (2007). *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, MA: MIT Press. ISBN: 9780262026192.

— (2015). *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, MA: The MIT Press. ISBN: 0262028565.

Borgman, Christine L. et al. (2007). "Drowning in Data: Digital Library Architecture to Support Scientific Use of Embedded Sensor Networks". In: *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries*. JCDL '07. Vancouver, BC, Canada: Association for Computing Machinery, pp. 269–277. DOI: 10.1145/1255175.1255228.

Botsman, Rachel (2017). *Who Can You trust? How Technology Brought Us Together - And Why It Could Drive Us Apart*. London, UK: Penguin Random House UK. ISBN: 9780241296172.

Bovens, Mark (2007). "Analysing and Assessing Accountability: A Conceptual Framework". In: *European Law Journal* 13.4, pp. 447–468. DOI: 10.1111/j.1468-0386.2007.00378.x.

Bowker, Geoffrey C. (2000). "Biodiversity Datadiversity". In: *Social Studies of Science* 30.5, pp. 643–683. DOI: 10.1177/030631200030005001.

Bowker, Geoffrey C. and Susan Leigh Star (2000). *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press. ISBN: 9780262522953.

Bradner, Erin (2001). "Social Affordances of Computer-Mediated Communication Technology: Understanding Adoption". In: *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, pp. 67–68. DOI: 10.1145/634067.634111.

Bradner, Erin, Wendy A. Kellogg, and Thomas Erickson (1999). "The Adoption and Use of "BABBLE": A Field Study of Chat in the Workplace". In: *Proceedings of the Sixth European Conference on Computer Supported Cooperative Work*. Copenghagen, Denmark, pp. 139–158. DOI: 10.1007/978-94-011-4441-4_8.

Bradner, Erin and Gloria Mark (2002). "Why Distance Matters: Effects on Cooperation, Persuasion and Deception". In: *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*. CSCW '02. New Orleans, Louisiana, USA: Association for Computing Machinery, pp. 226–235. DOI: 10.1145/587078.587110.

Braun, Virginia and Victoria Clarke (2006). "Using thematic analysis in psychology". In: *Qualitative Research in Psychology* 3.2, pp. 77–101. DOI: 10.1191/1478088706qp063oa.

Bridle, Helen et al. (2013). "Preparing for an Interdisciplinary Future: A Perspective from Early-career Researchers". In: *Futures* 53, pp. 22–32. DOI: 10.1016/j.futures.2013.09.003.

Brookbanks, Mike and Glenn Parry (2022). "The Impact of a Blockchain Platform on Trust in Established Relationships: A Case Study of Wine Supply Chains". In: *Supply Chain Management: An International Journal* 27 (7), pp. 128–146. DOI: 10.1108/SCM-05-2021-0227.

Brown, John Seely and Paul Duguid (2001). "Knowledge and Organization: A Social-Practice Perspective". In: *Organization Science* 12.2, pp. 198–213. DOI: 10.1287/orsc.12.2.198.10116.

Bryman, Alan (2008). *Social Research Methods*. 3rd ed. Oxford, UK: Oxford University Press. ISBN: 9780199202959.

Button, Graham (2006). "Introduction: A New Perspective on the Dependability of Software Systems". In: *Trust in Technology: A Socio-Technical Perspective*. Ed. by Karen Clarke et al. Dordrecht, NL: Springer Netherlands, pp. ix–xxv. ISBN: 9781402042584.

Calder, Muffy et al. (2018). "Making Sense of the World: Framing Models for Trustworthy Sensor-Driven Systems". In: *Computers* 7.4, p. 62. DOI: 10.3390/computers7040062.

Candela, Leonardo, Donatella Castelli, and Pasquale Pagano (2013). "Virtual Research Environments: An Overview and a Research Agenda". In: *Data Science Journal* 12, pp. 75–81. DOI: 10.2481/dsj.GRDI-013.

Carlson, Samuelle and Ben Anderson (2007). "What Are Data? The Many Kinds of Data and Their Implications for Data Re-Use". In: *Journal of Computer-Mediated Communication* 12.2, pp. 635–651. DOI: 10.1111/j.1083-6101.2007.00342.x.

Carroll, John M. (1996). "Encountering Others: Reciprocal Openings in Participatory Design and User-Centered Design". In: *Human Computer Interaction* 11.3, pp. 285–290. DOI: 10.1207/s15327051hci1103_5.

Castaldo, Sandro, Katia Premazzi, and Fabrizio Zerbini (2010). "The Meaning(s) of Trust: A Content Analysis on the Diverse Conceptualizations of Trust in Scholarly Research on Business Relationships". In: *Journal of Business Ethics* 96.4, pp. 657–668. DOI: 10.1007/s10551-010-0491-4.

Castelfranchi, Cristiano and Rino Falcone (2010). "Definitions of Trust: From Conceptual Components to the General Core". In: *Trust Theory: A Socio-*

Cognitive and Computational Model. John Wiley & Sons, Ltd. Chap. 1, pp. 7–33. ISBN: 9780470519851. DOI: 10.1002/9780470519851.ch1.

Castelfranchi, Cristiano and Rino Falcone (2016). "Trust & Self-Organising Socio-technical Systems". In: *Trustworthy Open Self-Organising Systems*. Ed. by Wolfgang Reif et al. Autonomic Systems. Springer. Chap. 8, pp. 209–229. DOI: 10.1007/978-3-319-29201-4_8.

— (2020). "Trust: Perspectives in Cognitive Science". In: *The Routledge Handbook of Trust and Philosophy*. Ed. by Judith Simon. New York, NY: Routledge. Chap. 17, pp. 214–228. ISBN: 9781315542294.

Charmaz, Kathy (2009). "Shifting the Grounds: Constructivist Grounded Theory Methods". In: *Developing Grounded Theory: The Second Generation*. Ed. by Janice M. Morse et al. Walnut Creek, CA: SAGE Publications, pp. 127–155. ISBN: 9781315430577.

Cheney, James, Laura Chiticariu, and Wang-Chiew Tan (2009). "Provenance in Databases: Why, How, and Where". In: *Found. Trends Databases* 1.4, pp. 379–474. DOI: 10.1561/1900000006.

Chin, George and Carina S. Lansing (2004). "Capturing and Supporting Contexts for Scientific Data Sharing via the Biological Sciences Collaboratory". In: *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*. CSCW '04. Chicago, Illinois, USA: Association for Computing Machinery, pp. 409–418. ISBN: 1581138105. DOI: 10.1145/1031607.1031677.

Clark, Charlotte et al. (2007). "The Messy Process of Research: Dilemmas, Process, and Critique". In: *Canadian Journal of Environmental Education* 12. Accessed 13/10/2020, pp. 110–126. URL: https://cjee.lakeheadu.ca/article/view/625.

Clark, David (2014). "The Role of Trust in Cyberspace". In: *Trust, Computing, and Society*. Ed. by Richard H.R. Harper. Cambridge: Cambridge University Press. Chap. 2, pp. 17–37. DOI: 10.1017/CBO9781139828567.

Clarke, Karen et al. (2006). "Trust and Organisational Work". In: *Trust in Technology: A Socio-Technical Perspective*. Ed. by Karen Clarke et al. Dordrecht, NL: Springer Netherlands. Chap. 1, pp. 1–20. ISBN: 9781402042584.

Clément, Fabrice (2020). "Trust: Perspectives in Psychology". In: *The Routledge Handbook of Trust and Philosophy*. Ed. by Judith Simon. New York, NY: Routledge. Chap. 16, pp. 205–213. ISBN: 9781315542294.

Cofta, Piotr (2007). "Confidence, Trust and Identity". In: *BT Technology Journal* 25.2, pp. 173–178. DOI: 10.1007/s10550-007-0042-4.

— (2021). "Trust, Choice, and Self-Preservation: A Computational Approach". In: *Cognition, Technology & Work* 23.1, pp. 85–101. DOI: 10.1007/s10111-019-00600-4.

Coleman, James S. (1990). *Foundations of Social Theory*. Cambridge, MA: Harvard University Press. ISBN: 0674312252.

Collins, Harry M. (1992). *Changing Order: Replication and Induction in Scientific Practice*. Chicago, IL: University of Chicago Press. ISBN: 0226113760.

Cook, Karen S. and Jessica J. Santana (2020). "Trust: Perspectives in Sociology". In: *The Routledge Handbook of Trust and Philosophy*. Ed. by Judith Simon. New York, NY: Routledge. Chap. 12, pp. 189–204. ISBN: 9781315542294.

Corritore, Cynthia L., Beverly Kracher, and Susan Wiedenbeck (2003). "On-line Trust: Concepts, Evolving Themes, A Model". In: *International Journal of Human-Computer Studies* 58.6, pp. 737–758. DOI: 10.1016/S1071-5819(03)00041-7.

Costa, Ana Cristina, C. Ashley Fulmer, and Neil R. Anderson (2018). "Trust in Work Teams: An Integrative Review, Multilevel Model, and Future Directions". In: *Journal of Organizational Behavior* 39.2, pp. 169–184. DOI: https://doi.org/10.1002/job.2213.

Costantino, Tracie E. (2008). "Constructivism". In: *The SAGE Encyclopedia of Qualitative Research Methods*. Ed. by Lisa M. Given. Thousand Oaks, CA: SAGE Publications, pp. 116–119. ISBN: 9781412941631.

Crabtree, Andy, David M. Nichols, et al. (2000). "Ethnomethodologically Informed Ethnography and Information System Design". In: *Journal of the American Society for Information Science* 51.7, pp. 666–682. URL: https://asistdl.onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-4571(2000)51:7%3C666::AID-ASI8%3E3.0.CO;2-5.

Crabtree, Benjamin F. and William L. Miller (1999). *Doing Qualitative Research*. 2nd ed. Thousand Oaks, CA: SAGE Publications. ISBN: 9780761914983.

Craggs, Barnaby and Awais Rashid (2019). "Trust Beyond Computation Alone: Human Aspects of Trust in Blockchain Technologies". In: *Proceedings of the 41st International Conference on Software Engineering: Software Engineering in Society*. ICSE-SEIS '19. Montréal, Québec, Canada, pp. 21–30. DOI: 10.1109/ICSE-SEIS.2019.00011.

Cragin, Melissa H. and Kalpana Shankar (2006). "Scientific Data Collections and Distributed Collective Practice". In: *Computer Supported Cooperative Work* 15, pp. 185–204. DOI: 10.1007/s10606-006-9018-z.

Crang, Mike and Ian Cook (2007). *Doing Ethnographies*. SAGE Research Methods. London: SAGE Publications. DOI: 10.4135/9781849208949.

Creswell, John W. (2013). *Qualitative Inquiry & Research Design*. 3rd ed. Thousand Oaks, CA: SAGE Publications. ISBN: 9781412995313.

Creswell, John W. and Dana L. Miller (2000). "Determining Validity in Qualitative Inquiry". In: *Theory Into Practice* 39.3, pp. 124–130. URL: http://www.jstor.org/stable/1477543.

Crotty, Michael (1998). *The Foundations of Social Research Meaning and Perspective in the Research Process*. London: SAGE Publications. ISBN: 9780761961062.

D'Cruz, Jason (2018). "Trust Within Limits". In: *International Journal of Philosophical Studies* 26.2, pp. 240–250. DOI: 10.1080/09672559.2018.1450080.

— (2019). "Humble Trust". In: *Philosophical Studies* 176.4 (4), pp. 933–953. DOI: 10.1007/s11098-018-1220-6.

— (2020). "Trust and Distrust". In: *The Routledge Handbook of Trust and Philosophy*. Ed. by Judith Simon. New York, NY: Routledge. Chap. 3, pp. 41–51. ISBN: 9781315542294.

Davis, Christine S. (2008). "Representation". In: *The SAGE Encyclopedia of Qualitative Research Methods*. Ed. by Lisa M. Given. Thousand Oaks, CA: SAGE Publications, p. 756. ISBN: 9781412941631.

Davis, Jenny L. and James B. Chouinard (2016). "Theorizing Affordances: From Request to Refuse". In: *Bulletin of Science, Technology & Society* 36.4, pp. 241–248. DOI: 10.1177/0270467617714944.

Davis, Matthew C., Rose Challenger, et al. (2014). "Advancing Socio-Technical Systems Thinking: A Call for Bravery". In: *Applied Ergonomics* 45.2, Part A, pp. 171–180. DOI: https://doi.org/10.1016/j.apergo.2013.02.009.

De Filippi, Primavera, Morshed Mannan, and Wessel Reijers (2020). "Blockchain as a Confidence Machine: The Problem of Trust & Challenges of Governance". In: *Technology in Society* 62, pp. 1–14. DOI: https://doi.org/10.1016/j.techsoc.2020.101284.

de Lusignan, Simon et al. (2011). "Key Concepts to Assess the Readiness of Data for International Research: Data Quality, Lineage and Provenance, Extraction and Processing Errors, Traceability, and Curation". In: *Yearbook of Medical Informatics* 6, pp. 112–120. DOI: 10.1055/s-0038-1638748.

De Roure, David, Carole Goble, Sergejs Aleksejevs, et al. (2010). "Towards Open Science: The MyExperiment Approach". In: *Concurrency and Computation: Practice and Experience* 22.17, pp. 2335–2353. DOI: https://doi.org/10.1002/cpe.1601.

De Roure, David, Carole Goble, Jiten Bhagat, et al. (2008). "myExperiment: Defining the Social Virtual Research Environment". In: *2008 IEEE Fourth International Conference on eScience*, pp. 182–189. DOI: 10.1109/eScience.2008.86.

Dean, Graham et al. (2017). "An Interdisciplinary approach to Finding and Using Data for Complex Environmental Modelling Problems: A Soil System Example". In: *EnviroInfo2017*. Accessed 06/09/2020. URL: http://www.research.lancs.ac.uk/portal/en/publications/an-interdisciplinary-approach-to-finding-and-using-data-for-complex-environmental-modelling-problems(c7bc7483-662b-4817-bd0b-5520fb970948).html.

Denning, Tamara et al. (2010). "Patients, Pacemakers, and Implantable Defibrillators: Human Values and Security for Wireless Implantable Medical Devices". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, pp. 917–926. DOI: 10.1145/1753326.1753462.

Denscombe, Martyn (1998). *The Good Research Guide: For Small-Scale Social Research Projects*. Buckingham: Open University Press. ISBN: 0335198058.

Denzin, Norman K. (1989). *Interpretive Interactionism*. Newbury Park, CA: Sage Publications. ISBN: 9780761915140.

Dombi, Judit, Tetyana Sydorenko, and Veronika Timpe-Laughlin (2022). "Common Ground, Cooperation, and Recipient Design in Human-Computer Interactions". In: *Journal of Pragmatics* 193, pp. 4–20. DOI: https://doi.org/10.1016/j.pragma.2022.03.001.

Donaldson, Devan Ray and Paul Conway (2015). "User Conceptions of Trustworthiness for Digital Archival Documents". In: *Journal of the Association for Information Science and Technology* 66.12, pp. 2427–2444. DOI: 10.1002/asi.23330.

Dourish, Paul (2001). *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge, MA: The MIT Press. ISBN: 9780262256056. DOI: 10.7551/mitpress/7221.001.0001.

— (2006). "Implications for Design". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '06. Montréal, Québec, Canada:

Association for Computing Machinery, pp. 541–550. ISBN: 1595933727. DOI: 10.1145/1124772.1124855.

Dourish, Paul and Edgar Gómez Cruz (2018). "Datafication and Data Fiction: Narrating Data and Narrating with Data". In: *Big Data & Society* 5.2, pp. 1–10. DOI: 10.1177/2053951718784083.

Douthit, Brian J. et al. (2021). "A Conceptual Framework of Data Readiness: The Contextual Intersection of Quality, Availability, Interoperability, and Provenance". In: *Applied Clinical Informatics* 12.3, pp. 675–685. DOI: 10.1055/s-0041-1732423.

Dwyer, Natasha, Anirban Basu, and Stephen Marsh (2013). "Reflections on Measuring the Trust Empowerment Potential of a Digital Environment". In: *Trust Management VII*. Ed. by M. Carmen Fernández Gago et al. Vol. 401. IFIP Advances in Information and Communication Technology. Springer, pp. 127–135. URL: https://doi.org/10.1007/978-3-642-38323-6_9.

Dwyer, Natasha and Stephen Marsh (2016). "To Trust or Distrust: Has a Digital Environment Empowered Users to Proceed on Their Own Terms?" In: *Trustworthy Open Self-Organising Systems*. Ed. by Wolfgang Reif et al. Cham: Springer International Publishing. Chap. 9, pp. 231–244. ISBN: 978-3-319-29201-4. DOI: 10.1007/978-3-319-29201-4_9.

Easterbrook, Steve M. (2010). "Climate Change: A Grand Software Challenge". In: *Proceedings of the Workshop on Future of Software Engineering Research, FoSER 2010, at the 18th ACM SIGSOFT International Symposium on Foundations of Software Engineering, 2010, Santa Fe, NM, USA, November 7-11, 2010*. Ed. by Gruia-Catalin Roman and Kevin J. Sullivan. ACM, pp. 99–104. DOI: 10.1145/1882362.1882383.

— (2014). "From Computational Thinking to Systems Thinking: A conceptual toolkit for sustainability computing". In: *ICT for Sustainability 2014 (ICT4S-14), Stockholm, Sweden, August 25, 2014*. Atlantis Press. DOI: 10.2991/ict4s-14.2014.28.

Edwards, Paul N. (2010). *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge, MA: The MIT Press. ISBN: 0262013924.

Edwards, Paul N. et al. (2011). "Science Friction: Data, Metadata, and Collaboration". In: *Social Studies of Science* 41.5, pp. 667–690. DOI: 10.1177/0306312711413314.

Ehsan, Upol et al. (2021). "Expanding Explainability: Towards Social Transparency in AI Systems". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. New York, NY: Association for Computing Machinery, pp. 1–19. DOI: 10.1145/3411764.3445188.

Elkhatib, Yehia, Gordon S. Blair, and Bholanathsingh Surajbali (2013). "Experiences of Using a Hybrid Cloud to Construct an Environmental Virtual Observatory". In: *Proceedings of the 3rd International Workshop on Cloud Data and Platforms, CloudDP@EuroSys 2013, Prague, Czech Republic, April 14-17, 2013*. Ed. by Paulo Ferreira, Zbigniew Jerzak, and Peter R. Pietzuch. ACM, pp. 13–18. DOI: 10.1145/2460756.2460759.

Elkhatib, Yehia, Alastair L. Gemmell, et al. (2019). "Widening the Circle of Engagement Around Environmental Issues using Cloud-based Tools". In: *39th IEEE International Conference on Distributed Computing Systems, ICDCS 2019, Dallas, TX, USA, July 7-10, 2019*. IEEE, pp. 1404–1415. DOI: 10.1109/ICDCS.2019.00140.

Elsden, Chris et al. (2018). "Making Sense of Blockchain Applications: A Typology for HCI". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. New York, NY, USA: Association for Computing Machinery, pp. 1–14. DOI: 10.1145/3173574.3174032.

Erickson, Thomas (1996). "Design as Storytelling". In: *Interactions* 3.4, pp. 30–35. DOI: 10.1145/234813.234817.

Erickson, Thomas and Wendy A. Kellogg (2000). "Social Translucence: An Approach to Designing Systems That Support Social Processes". In: *ACM Transactions on Computer-Human Interaction* 7.1, pp. 59–83. DOI: 10.1145/344949.345004.

Eriksson, Per Erik and Yvonne Eriksson (2019). "Live-action Communication Design: A Technical How-To Video Case Study". In: *Technical Communication Quarterly* 28.1, pp. 69–91. DOI: 10.1080/10572252.2018.1528388.

Ess, Charles M. (2020). "Trust and Information and Communication Technologies". In: *The Routledge Handbook of Trust and Philosophy*. Ed. by Judith Simon. New York, NY: Routledge. Chap. 31, pp. 405–420. ISBN: 9781315542294.

European Commission (2015). *e-Infrastructures for Virtual Research Environments (VRE)*. Accessed 18/12/2021. URL: https://cordis.europa.eu/programme/id/H2020_EINFRA-9-2015.

Evans, Sandra K. et al. (2016). "Explicating Affordances: a Conceptual Framework for Understanding Affordances in Communication Research". In: *Journal of Computer-Mediated Communication* 22.1, pp. 35–52. DOI: 10.1111/jcc4.12180.

Faniel, Ixchel M., Rebecca D. Frank, and Elizabeth Yakel (2019). "Context from the Data Reuser's Point of View". In: *Journal of Documentation* 75 (6), pp. 1274–1297. DOI: 10.1108/JD-08-2018-0133.

Faniel, Ixchel M. and Trond E. Jacobsen (2010). "Reusing Scientific Data: How Earthquake Engineering Researchers Assess the Reusability of Colleagues' Data". In: *Computer Supported Cooperative Work* 3 (19), pp. 355–375. DOI: 10.1007/s10606-010-9117-8.

Faniel, Ixchel M., Adam Kriesberg, and Elizabeth Yakel (2016). "Social Scientists' Satisfaction with Data Reuse". In: *Journal of the Association for Information Science and Technology* 67 (6), pp. 1404–1416. DOI: 10.1002/asi.23480.

Faniel, Ixchel M. and Elizabeth Yakel (2017). "Practices Do Not Make Perfect: Disciplinary Data Sharing and Reuse Practices and their Implications for Repository Data Curation". In: *Curating Research Data Volume 1: Practical Strategies for Your Digital Repository*. Ed. by Lisa R. Johnston. Chicago, IL: Association of College and Research Libraries Press. Chap. 4, pp. 103–126. ISBN: 9780838988589.

Faniel Ixchel, M. and Ann Zimmerman (2011). "Beyond the Data Deluge: A Research Agenda for Large-Scale Data Sharing and Reuse". In: *International Journal of Digital Curation* 6 (1), pp. 58–69. URL: https://ijdc.net/index.php/ijdc/article/view/172.

Faraj, Samer and Bijan Azad (2012). "The Materiality of Technology: An Affordance Perspective". In: *Materiality and Organizing*. Ed. by Paul M. Leonardi, Bonnie A. Nardi, and Jannis Kallinikos. Oxford: Oxford University Press, pp. 237–258. ISBN: 9780199664054.

Fayard, Anne-Laure and John Weeks (2014). "Affordances for Practice". In: *Information and Organisation* 24.4, pp. 236–249. DOI: 10.1016/j.infoandorg.2014.10.001.

Fear, Kathleen and Devan Ray Donaldson (2012). "Provenance and Credibility in Scientific Data Repositories". In: *Archival Science* 12.3, pp. 319–339. DOI: 10.1007/s10502-012-9172-7.

Feinberg, Melanie (2017). "A Design Perspective on Data". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17, pp. 2952–2963. DOI: 10.1145/3025453.3025837.

Finholt, Thomas A. and Gary M. Olson (1997). "From Laboratories to Collaboratories: A New Organizational Form for Scientific Collaboration". In: *Psychological Science* 8.1, pp. 28–36. DOI: 10.1111/j.1467-9280.1997.tb00540.x.

Flechais, Ivan, Jens Riegelsberger, and M. Angela Sasse (2005). "Divide and Conquer: The Role of Trust and Assurance in the Design of Secure Socio-Technical Systems". In: *Proceedings of the 2005 Workshop on New Security Paradigms*. NSPW '05. Lake Arrowhead, California: Association for Computing Machinery, pp. 33–41. ISBN: 1595933174. DOI: 10.1145/1146269.1146280.

Frost-Arnold, Karen (2020). "Trust and Epistemic Responsibility". In: *The Routledge Handbook of Trust and Philosophy*. Ed. by Judith Simon. New York, NY: Routledge. Chap. 5, pp. 64–75. ISBN: 9781315542294.

Fukuyama, Francis (1995). *Trust: The Social Virtues and the Creation of Prosperity*. New York City, NY: Free Press. ISBN: 9780684825250.

Fulmer, Ashley and Kurt Dirks (2018). "Multilevel Trust: A Theoretical and Practical Imperative". In: *Journal of Trust Research* 8.2, pp. 137–141. DOI: 10.1080/21515581.2018.1531657.

Fusch, Patricia I. and Lawrence R. Ness (2015). "Are We There Yet? Data Saturation in Qualitative Research". In: *The Qualitative Report* 20.9, pp. 1408–1416. DOI: 10.46743/2160-3715/2015.2281.

Gach, Katie Z. and Jed R. Brubaker (2020). "Experiences of Trust in Postmortem Profile Management". In: *ACM Transactions on Social Computing* 3.1. DOI: 10.1145/3365525.

Gambetta, Diego (1988). *Trust: Making and Breaking of Cooperative Relations*. Oxford, UK: Basil Blackwell. ISBN: 0631155066.

Gamble, Matthew and Carole Goble (2011). "Quality, Trust, and Utility of Scientific Data on the Web: Towards a Joint Model". In: *Proceedings of the 3rd International Web Science Conference*. Association for Computing Machinery. DOI: 10.1145/2527031.2527048.

Gaver, Bill and John Bowers (2012). "Annotated Portfolios". In: *Interactions* 19.4, pp. 40–49. DOI: 10.1145/2212877.2212889.

Gaver, William W. (1991). "Technology Affordances". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 79–84. DOI: 10.1145/108844.108856.

— (1992). "The Affordances of Media Spaces for Collaboration". In: *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work*, pp. 17–24. DOI: 10.1145/143457.371596.

Gebru, Timnit et al. (2021). "Datasheets for Datasets". In: *Commun. ACM* 64.12, pp. 86–92. DOI: 10.1145/3458723.

Geertz, Clifford (1973). *The Interpretation of Cultures*. New York: Basic Books. ISBN: 046503425X.

Gefen, David, Izak Benbasat, and Paul A. Pavlou (2008). "A Research Agenda for Trust in Online Environments". In: *Journal of Management Information Systems* 24.4, pp. 275–286. ISSN: 07421222. URL: http://www.jstor.org/stable/40398920.

Gibson, James J. (1977). "The Theory of Affordances". In: *Perceiving, Acting, Knowing: Toward an Ecological Psychology*. Ed. by Robert Shaw and John Bransford. Hillsdale, NJ: Lawrence Erlbaum. Chap. 3, pp. 67–82. ISBN: 9780470990148.

— (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin. ISBN: 9781848725775.

Giddens, Anthony (1990). *The Consequences of Modernity*. Cambridge, UK: Polity Press. ISBN: 0745609236.

Gordov, Evgeny P. et al. (2018). "A Virtual Thematic Research Environment for Analysis of Climate Change and its Consequences for Northern Eurasia: The Current State of the Art and Perspectives". In: *IOP Conference Series: Earth and Environmental Science* 211. DOI: 10.1088/1755-1315/211/1/012079.

Grudin, Jonathan (1988). "Why CSCW Applications Fail: Problems in the Design and Evaluation of Organizational Interfaces". In: *Proceedings of the 1988 ACM Conference on Computer-Supported Cooperative Work*. CSCW '88. Portland, Oregon, USA: Association for Computing Machinery, pp. 85–93. DOI: 10.1145/62266.62273.

— (1994a). "Computer-Supported Cooperative Work: History and Focus". In: *Computer* 27.5, pp. 19–26. DOI: 10.1109/2.291294.

— (1994b). "Groupware and Social Dynamics: Eight Challenges for Developers". In: *Commun. ACM* 37.1, pp. 92–105. DOI: 10.1145/175222.175230.

Guba, Egon G. and Yvonna S. Lincoln (1994). "Competing Paradigms in Qualitative Research". In: *Handbook of Qualitative Research*. Ed. by Norman K. Denzin and Yvonna S. Lincoln. Thousand Oaks, CA: SAGE Publications. Chap. 6, pp. 105–117. ISBN: 9780803946798.

Gulati, Siddharth, Sonia Sousa, and David Lamas (2017). "Modelling Trust: An Empirical Assessment". In: *Human-Computer Interaction – INTERACT 2017*. Ed. by Regina Bernhaupt et al., pp. 40–61. DOI: 10.1007/978-3-319-68059-0_3.

— (2019). "Design, Development and Evaluation of a Human-Computer Trust Scale". In: *Behaviour & Information Technology* 38.10, pp. 1004–1015. DOI: 10.1080/0144929X.2019.1656779.

Hallgren, Willow et al. (2016). "The Biodiversity and Climate Change Virtual Laboratory: Where Ecology Meets Big Data". In: *Environmental Modelling & Software* 76, pp. 182–186. ISSN: 1364-8152. DOI: https://doi.org/10.1016/j.envsoft.2015.10.025.

Harper, Richard (2014). "Introduction and Overview". In: *Trust, Computing, and Society*. Ed. by Richard H.R. Harper. Cambridge: Cambridge University Press. Chap. 1, pp. 3–16. DOI: 10.1017/CBO9781139828567.

Harper, Richard H. R. (2000). "Analysing Work Practice and the Potential Role of New Technology at the International Monetary Fund: Some Remarks on the Role of Ethnomethodology". In: *Workplace Studies: Recovering Work Practice and Informing System Design*. Ed. by Paul Luff, Jon Hindmarsh, and Christian Heath. Cambridge, UK: Cambridge University Press. Chap. 8, pp. 169–186. ISBN: 0521598214.

Hartson, Rex (2003). "Cognitive, Physical, Sensory, and Functional Affordances in Interaction Design". In: *Behaviour & Information Technology* 22.5, pp. 315–338. DOI: 10.1080/01449290310001592587.

Harz, Dominik and Magnus Boman (2019). "The Scalability of Trustless Trust". In: *Financial Cryptography and Data Security*. Ed. by Aviv Zohar et al. Springer Berlin Heidelberg, pp. 279–293. DOI: 10.1007/978-3-662-58820-8_19.

Hays, Samuel P. (2000). *A History of Environmental Politics Since 1945*. University of Pittsburgh Press. ISBN: 9780822941286. URL: http://www.jstor.org/stable/j.ctt6wrcjm.

Henrys, Peter A. and Susan G. Jarvis (2019). "Integration of Ground Survey and Remote Sensing Derived Data: Producing Robust Indicators of Habitat Extent and Condition". In: *Ecology and Evolution* 9.14, pp. 8104–8112. DOI: 10.1002/ece3.5376.

Hertzum, Morten (1999). "Six Roles of Documents in Professionals' Work". In: *Proceedings of the Sixth European Conference on Computer Supported Cooperative Work*. Copenghagen, Denmark, pp. 41–60. DOI: 10.1007/978-94-011-4441-4_3.

Hertzum, Morten et al. (2002). "Trust in Information Sources: Seeking Information from People, Documents, and Virtual Agents". In: *Interacting with Computers* 14.5, pp. 575–599. DOI: 10.1016/S0953-5438(02)00023-1.

Hill, Raquel and Devan Ray Donaldson (2015). "Bridging the Trust Gap: Integrating Models of Behavior and Perception". In: *Proceedings of the 2015 New Security Paradigms Workshop*. NSPW '15. Twente, Netherlands: Association for Computing Machinery, pp. 148–155. DOI: 10.1145/2841113.2841125.

Hind, Michael et al. (2020). "Experiences with Improving the Transparency of AI Models and Services". In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI EA '20. Honolulu, HI, USA: Association for Computing Machinery, pp. 1–8. DOI: 10.1145/3334480.3383051.

Hine, Christine (2006). "Databases as Scientific Instruments and Their Role in the Ordering of Scientific Work". In: *Social Studies of Science* 36.2, pp. 269–298. DOI: 10.1177/0306312706054047.

Hoehle, Hartmut, Sid Huff, and Sigi Goode (2012). "The Role of Continuous Trust in Information Systems Continuance". In: *Journal of Computer Information Systems* 52.4, pp. 1–9. URL: https://www.tandfonline.com/doi/abs/10.1080/08874417.2012.11645571.

Hoffman, Michał R. et al. (2018). "Smart Papers: Dynamic Publications on the Blockchain". In: *The Semantic Web*. Ed. by Aldo Gangemi et al., pp. 304–318. DOI: 10.1007/978-3-319-93417-4_20.

Hollan, Jim and Scott Stornetta (1992). "Beyond Being There". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '92. New York, NY: Association for Computing Machinery, pp. 119–125. ISBN: 0897915135. DOI: 10.1145/142750.142769.

Holland, Sarah et al. (2018). *The Dataset Nutrition Label: A Framework To Drive Higher Data Quality Standards*. arXiv: 1805.03677. URL: http://datanutrition.media.mit.edu/abstract.pdf.

Hollaway, Michael et al. (2020). "Tackling the Challenges of 21st Open Science and Beyond: A Data Science Lab Approach". In: *Patterns*. DOI: https://doi.org/10.1016/j.patter.2020.100103.

Holtham, Clive (1995). "Integrating Technologies to Support Action". In: *Interacting with Computers* 7.1, pp. 91–107. ISSN: 0953-5438. DOI: 10.1016/0953-5438(95)90821-2.

Holzer, Jennifer M. et al. (2019). "Evaluating Transdisciplinary Science to Open Research: Implementation Spaces in European Social-ecological Systems". In: *Biological Conservation* 238, p. 108228. ISSN: 0006-3207. DOI: https://doi.org/10.1016/j.biocon.2019.108228.

Hunhevicz, Jens J. and Daniel M. Hall (2020). "Do You Need A Blockchain in Construction? Use Case Categories and Decision Framework for DLT Design Options". In: *Advanced Engineering Informatics* 45, pp. 1–14. DOI: https://doi.org/10.1016/j.aei.2020.101094.

Hutchby, Ian (2001). "Technologies, Texts and Affordances". In: *Sociology* 35.2, pp. 441–456. DOI: 10.1177/S0038038501000219.

Hyett, Nerida, Amanda Kenny, and Virginia Dickson-Swift (2014). "Methodology or Method? A Critical Review of Qualitative Case Study Reports". In: *International*

Journal of Qualitative Studies on Health and Well-Being 9.1. DOI: 10.3402/qhw.v9.23606.

Islind, Anna Sigridur et al. (2019). "The Virtual Clinic: Two-Sided Affordances in Consultation Practice". In: *Computer Supported Cooperative Work* 28.3–4, pp. 435–468. DOI: 10.1007/s10606-019-09350-3.

Jacovi, Alon et al. (2021). "Formalizing Trust in Artificial Intelligence: Prerequisites, Causes and Goals of Human Trust in AI". In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '21. Virtual Event, Canada: Association for Computing Machinery, pp. 624–635. DOI: 10.1145/3442188.3445923.

Jantz, Ronald and Michael Giarlo (2007). "Digital Archiving and Preservation: Technologies and Processes for a Trusted Repository". In: *Journal of Archival Organization* 4.1-2, pp. 193–213. URL: https://doi.org/10.1300/J201v04n01_10.

Jeffery, Keith, Leonardo Candela, and Helen Glaves (2020). "Virtual Research Environments for Environmental and Earth Sciences: Approaches and Experiences". In: *Towards Interoperable Research Infrastructures for Environmental and Earth Sciences: A Reference Model Guided Approach for Common Challenges*. Ed. by Zhiming Zhao and Margareta Hellström. Cham: Springer International Publishing, pp. 272–289. ISBN: 978-3-030-52829-4. DOI: 10.1007/978-3-030-52829-4_15.

Jirotka, Marina et al. (2005). "Collaboration and Trust in Healthcare Innovation: The EDiaMoND Case Study". In: *Computer Supported Cooperative Work* 14.4, pp. 369–398. ISSN: 0925-9724. DOI: 10.1007/s10606-005-9001-0.

Jones, Gareth R. and Jennifer M. George (1998). "The Experience and Evolution of Trust: Implications for Cooperation and Teamwork". In: *The Academy of Management Review* 23.3, pp. 531–546. DOI: 10.2307/259293.

Jones, Karen (1996). "Trust as an Affective Attitude". In: *Ethics* 107.1, pp. 4–25. URL: https://www.jstor.org/stable/2382241.

— (2019). "Trust, Distrust, and Affective Looping". In: *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 176.4, pp. 955–968. DOI: 10.1007/s11098-018-1221-5.

Kannengiesser, Udo and John S. Gero (2012). "A Process Framework of Affordances in Design". In: *Design Issues* 28.1, pp. 50–62. DOI: 10.1162/DESI_a_00123.

Kaplan, Simon and Lesley Seebeck (2001). "Harnessing Complexity in CSCW". In: *Proceedings of the Seventh Conference on European Conference on Computer Supported Cooperative Work*. ECSCW'01. Bonn, Germany: Kluwer Academic Publishers, pp. 359–378. ISBN: 0792371623. URL: https://link.springer.com/chapter/10.1007/0-306-48019-0_19.

Kaptelinin, Victor and Bonnie Nardi (2012). "Affordances in HCI: Toward a Mediated Action Perspective". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Austin, Texas, USA: Association for Computing Machinery, pp. 967–976. DOI: 10.1145/2207676.2208541.

Kay, Judy and Bob Kummerfeld (2010). "Tackling HCI Challenges of Creating Personalised, Pervasive Learning Ecosystems". In: *Sustaining TEL: From Innovation to Learning and Practice*. Ed. by Martin Wolpers et al. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–16. URL: https://doi.org/10.1007/978-3-642-16020-2_1.

Kelton, Kari, Kenneth R. Fleischmann, and William A. Wallace (2008). "Trust in Digital Information". In: *Journal of the American Society for Information Science and Technology* 59.3, pp. 363–374. DOI: https://doi.org/10.1002/asi.20722.

Kiani, Kimia et al. (2019). "Beyond "One-Size-Fits-All": Understanding the Diversity in How Software Newcomers Discover and Make Use of Help Resources". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19, pp. 1–14. DOI: 10.1145/3290605.3300570.

Kivunja, Charles and Ahmed Bawa Kuyini (2017). "Understanding and Applying Research Paradigms in Educational Contexts". In: *International Journal of*

Higher Education 6.5. Accessed 18/11/2020, pp. 1–26. URL: https://ideas.repec.org/a/jfr/ijhe11/v6y2017i5p26.html.

Knorr Cetina, Karin (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press. ISBN: 9780674258945.

Knowles, Bran (2016). "Emerging Trust Implications of Data-Rich Systems". In: *IEEE Pervasive Computing* 15.4, pp. 76–84. DOI: 10.1109/MPRV.2016.68.

Knowles, Bran, Mike Harding, et al. (2014). "Trustworthy by Design". In: *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. CSCW '14. Baltimore, Maryland, USA: Association for Computing Machinery, pp. 1060–1071. ISBN: 9781450325400. DOI: 10.1145/2531602.2531699.

Knowles, Bran, Mark Rouncefield, et al. (2015). "Models and Patterns of Trust". In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. CSCW '15. Vancouver, BC, Canada: Association for Computing Machinery, pp. 328–338. ISBN: 9781450329224. DOI: 10.1145/2675133.2675154.

Knowles, Bran, Alison Smith-Renner, et al. (2018). "Uncertainty in Current and Future Health Wearables". In: *Communications of the ACM* 61.12, pp. 62–67. DOI: 10.1145/3199201.

Kobayashi, Guiou et al. (2015). "The Internet of Things and its Impact on Social Relationships Involving Mutual Trust". In: *2015 IEEE International Symposium on Technology and Society (ISTAS)*, pp. 1–6. DOI: 10.1109/ISTAS.2015.7439416.

Konrad, Kornelia, Gerhard Fuchs, and Jochen Barthel (1999). "Trust and Electronic Commerce More than a Technical Problem". In: *Proceedings of the 18th IEEE Symposium on Reliable Distributed Systems*. SRDS '99. USA: IEEE Computer Society, p. 360. DOI: 10.1109/RELDIS.1999.805124.

Kratz, John E. and Carly Strasser (2015). "Making Data Count". In: *Scientific Data* 2.1. DOI: 10.1038/sdata.2015.39.

Kroeger, Frens (2019). "Unlocking the Treasure Trove: How Can Luhmann's Theory of Trust Enrich Trust Research?" In: *Journal of Trust Research* 9.1, pp. 110–124. DOI: 10.1080/21515581.2018.1552592.

Kvale, Steinar (1996). *InterViews: An Introduction to Qualitative Research Interviewing*. Thousand Oaks, CA: Sage Publications. ISBN: 080395820X.

Lagerspetz, Olli (2014). "The Worry about Trust". In: *Trust, Computing, and Society*. Ed. by Richard H. Harper. Cambridge: Cambridge University Press. Chap. 6, pp. 120–143. DOI: 10.1017/CBO9781139828567.010.

Lahno, Bernd (2001a). "Institutional Trust: A Less Demanding Form of Trust?" In: *Revista Latinoamericana de Estudios Avanzados* 15, pp. 19–58.

— (2001b). "On the Emotional Character of Trust". In: *Ethical Theory and Moral Practice* 4, pp. 171–189. DOI: 10.1023/A:1011425102875.

— (2020). "Trust and Emotion". In: *The Routledge Handbook of Trust and Philosophy*. Ed. by Judith Simon. New York, NY: Routledge. Chap. 12, pp. 147–159. ISBN: 9781315542294.

Lankton, Nancy, D. Harrison Mcknight, and John Tripp (2015). "Technology, Humanness, and Trust: Rethinking Trust in Technology". In: *Journal of the Association for Information Systems* 16, pp. 880–918. DOI: 10.17705/1jais.00411.

Latour, Bruno and Steve Woolgar (1979). *Laboratory Life: The Construction of Scientific Facts*. Princeton: Princeton University Press. ISBN: 0691094187.

Lauckner, Heidi, Margo Paterson, and Terry Krupa (2012). "Using Constructivist Case Study Methodology to Understand Community Development Processes: Proposed Methodological Questions to Guide the Research Process". In: *The Qualitative Report* 17.13, pp. 1–22. DOI: 10.46743/2160-3715/2012.1790.

Law, John (2004). *After Method: Mess in Social Science Research*. London, UK: Routledge. ISBN: 0203481143.

Lee, Dong Joon and Besiki Stvilia (2017). "Practices of Research Data Curation in Institutional Repositories: A Qualitative View from Repository Staff". In: *PLOS ONE* 12.3, pp. 1–44. DOI: 10.1371/journal.pone.0173987.

Lee, John and Neville Moray (1992). "Trust, Control Strategies and Allocation of Function in Human-Machine Systems". In: *Ergonomics* 35.10, pp. 1243–1270. DOI: 10.1080/00140139208967392.

Lee, John D. and Katrina A. See (2004). "Trust in Automation: Designing for Appropriate Reliance". In: *Human Factors* 46.1, pp. 50–80. URL: https://doi.org/10.1518/hfes.46.1.50_30392.

Leeming, Dawn (2018). "The Use of Theory in Qualitative Research". In: *Journal of Human Lactation* 34.4, pp. 668–673. URL: https://journals.sagepub.com/doi/abs/10.1177/0890334418794666.

Lemieux, Victoria L. (2016). "Trusting Records: Is Blockchain Technology the Answer?" In: *Records Management Journal* 26.2, pp. 110–139. DOI: 10.1108/RMJ-12-2015-0042.

Leonardi, Paul M. (2011). "When Flexible Routines Meet Flexible Technologies: Affordance, Constraint, and the Imbrication of Human and Material Agencies". In: *MIS Quarterly* 35.1, pp. 147–167. DOI: 10.2307/23043493.

Levy, David M. (2000). "Where's Waldo? Reflections on Copies and Authenticity in a Digital Environment". In: *Authenticity in a Digital Environment*. Ed. by Charles T. Cullen et al. Council on Library and Information Resources. Chap. 3, pp. 24–31. ISBN: 1887334777.

Lewicki, Roy J. and Barbara Benedict Bunker (1996). "Developing and Maintaining Trust in Work Relationships". In: *Trust in Organizations: Frontiers of Theory and Research*. Ed. by Roderick M. Kramer and Tom Tyler. Thousand Oaks, CA: SAGE Publications, Inc. Chap. 7, pp. 114–139. DOI: 10.4135/9781452243610.n7.

Lewicki, Roy J., Daniel J. McAllister, and Robert J. Bies (1998). "Trust and Distrust: New Relationships and Realities". In: *The Academy of Management Review* 23.3, pp. 438–458. DOI: 10.2307/259288.

Lewis, J. David and Andrew J. Weigert (1985). "Trust as a Social Reality". In: *Social Forces* 63.4, pp. 967–985. DOI: 10.1093/sf/63.4.967.

— (2012). "The Social Dynamics of Trust: Theoretical and Empirical Research, (1985-2012)". In: *Social Forces* 91.1, pp. 25–31. ISSN: 0037-7732. DOI: 10.1093/sf/sos116.

Li, Hong Ling (2007). "From Shannon-Weaver to Boisot: A Review on the Research of Knowledge Transfer Model". In: *2007 International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 5439–5442. DOI: 10.1109/WICOM.2007.1332.

Li, Xin, Traci J. Hess, and Joseph S. Valacich (2008). "Why Do We Trust New Technology? A Study of Initial Trust Formation with Organizational Information Systems". In: *The Journal of Strategic Information Systems* 17.1, pp. 39–71. ISSN: 0963-8687. DOI: 10.1016/j.jsis.2008.01.001.

Liao, Q.Vera and S. Shyam Sundar (2022). "Designing for Responsible Trust in AI Systems: A Communication Perspective". In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '22. Seoul, Republic of Korea: Association for Computing Machinery, pp. 1257–1268. DOI: 10.1145/3531146.3533182.

Lin, Dawei et al. (2020). "The TRUST Principles for Digital Repositories". In: *Scientific Data* 7.144. DOI: 10.1038/s41597-020-0486-7.

Lincoln, Yvonna S. and Egon G. Guba (1985). *Naturalistic Inquiry*. Newbury Park, CA: SAGE Publications. ISBN: 9780803924314.

Lord, Philip et al. (2004). "From Data Deluge to Data Curation". In: *Proceedings of the UK E-Science All Hands Meeting*. Accessed 15/12/2020, pp. 371–375. URL: https://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/150.pdf.

Löwgren, Jonas (2013). "Annotated Portfolios and Other Forms of Intermediate-level Knowledge". In: *Interactions* 20.1, pp. 30–34. DOI: 10.1145/2405716.2405725.

Lub, Vasco (2015). "Validity in Qualitative Evaluation: Linking Purposes, Paradigms, and Perspectives". In: *International Journal of Qualitative Methods* 14.5, pp. 1–8. DOI: 10.1177/1609406915621406.

Lucas, Chloe, Peat Leith, and Aidan Davison (2015). "How Climate Change Research Undermines Trust in Everyday Life: A Review". In: *WIREs Climate Change* 6.1, pp. 79–91. DOI: 10.1002/wcc.320.

Luhmann, Niklas (2017). *Trust and Power*. Cambridge, UK: Polity Press. ISBN: 9781509519453.

Lyle, John and Andrew Martin (2010). "Trusted Computing and Provenance: Better Together". In: *Proceedings of the 2nd Conference on Theory and Practice of Provenance*. TAPP'10. Accessed 20/08/2020. San Jose, California: USENIX Association, p. 1. URL: https://www.usenix.org/legacy/events/tapp10/tech/full_papers/lyle.pdf.

Lynch, Clifford (2000). "Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust". In: *Authenticity in a Digital Environment*. Ed. by Charles T. Cullen et al. Council on Library and Information Resources. Chap. 4, pp. 32–50. ISBN: 1887334777.

Lyon, Fergus, Guido Möllering, and Mark N.K. Saunders (2015). "Researching Trust: The Ongoing Challenge of Matching Objectives and Methods". In: *Handbook of Research Methods on Trust*. Ed. by Fergus Lyon, Guido Möllering, and Mark N.K. Saunders. 2nd ed. Cheltenham, UK: Edward Elgar. Chap. 1, pp. 1–24. ISBN: 9781782547419. DOI: 10.4337/9781782547419.

Mackay, Wendy E. et al. (1998). "Reinventing the Familiar: Exploring an Augmented Reality Design Space for Air Traffic Control". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '98. Los Angeles, California, USA, pp. 558–565. DOI: 10.1145/274644.274719.

Maher, Carmel et al. (2018). "Ensuring Rigor in Qualitative Data Analysis: A Design Research Approach to Coding Combining NVivo With Traditional Material Methods". In: *International Journal of Qualitative Methods* 17.1, pp. 1–13. DOI: 10.1177/1609406918786362.

Mahr, Dominik and Jisu Huh (2022). "Technologies in Service Communication: Looking Forward". In: *Journal of Service Management* 33.4/5, pp. 648–656. DOI: 10.1108/JOSM-03-2022-0075.

Markus, M. Lynne (2001). "Toward a Theory of Knowledge Reuse: Types of Knowledge Reuse Situations and Factors in Reuse Success". In: *Journal of Management Information Systems* 18.1, pp. 57–93. URL: https://www.jstor.org/stable/40398517.

Marsh, Stephen, Tosan Atele-Williams, et al. (2020). "Thinking about Trust: People, Process, and Place". In: *Patterns* 1.3. DOI: https://doi.org/10.1016/j.patter.2020.100039.

Marsh, Stephen and Mark R. Dibben (2003). "The Role of Trust in Information Science and Technology". In: *Annual Review of Information Science and Technology* 37 (1), pp. 465–498. DOI: 10.1002/aris.1440370111.

Martin, Dave and Mark Rouncefield (2003). "Making the Organization Come Alive: Talking Through and About the Technology in Remote Banking". In: *Human–Computer Interaction* 18.1-2, pp. 111–148. URL: https://doi.org/10.1207/S15327051HCI1812_5.

Martin, David, Tom Rodden, et al. (2001). "Finding Patterns in the Fieldwork". In: *ECSCW 2001: Proceedings of the Seventh European Conference on Computer Supported Cooperative Work 16–20 September 2001, Bonn, Germany*. Ed. by Wolfgang Prinz et al. Springer Netherlands, pp. 39–58. DOI: 10.1007/0-306-48019-0_3.

Mason, Jennifer (2002). *Qualitative Researching*. 2nd ed. London: SAGE Publications. ISBN: 0761974288.

Mason, Mark (2010). "Sample Size and Saturation in PhD Studies Using Qualitative Interviews". In: *Forum: Qualitative Social Research* 11.3. DOI: 10.17169/fqs-11.3.1428.

Mattor, Katherine et al. (Oct. 2014). "Transdisciplinary Research on Environmental Governance: A View from the Inside". In: *Environmental Science & Policy* 42, pp. 90–100. DOI: 10.1016/j.envsci.2014.06.002.

Maxwell, Joseph A. (2009). "Designing a Qualitative Study". In: *The SAGE Handbook of Applied Social Research Methods*. Ed. by Leonard Bickman and Debra J. Rog. 2nd ed. Thousand Oaks, CA: SAGE Publications. Chap. 7, pp. 214–253. URL: https://dx.doi.org/10.4135/9781483348858.

Maxwell, Joseph A. and Kavita Mittapalli (2008). "Theory". In: *The SAGE Encyclopedia of Qualitative Research Methods*. Ed. by Lisa M. Given. Thousand Oaks, CA: SAGE Publications, pp. 876–879. ISBN: 9781412941631.

Mayer, Roger C., James H. Davis, and F. David Schoorman (1995). "An Integrative Model of Organizational Trust". In: *The Academy of Management Review* 20.3, pp. 709–734. URL: http://www.jstor.org/stable/258792.

Mazé, Ramia and Johan Redström (2005). "Form and the Computational Object". In: *Digital Creativity* 16.1, pp. 7–18. DOI: 10.1080/14626260500147736.

McEvily, Bill, Vincenzo Perrone, and Akbar Zaheer (2003). "Trust as an Organizing Principle". In: *Organization Science* 14.1, pp. 91–103. DOI: 10.1287/orsc.14.1.91.12814.

McGonigle, Daniel F. et al. (2020). "A Knowledge Brokering Framework for Integrated Landscape Management". In: *Frontiers in Sustainable Food Systems* 4. DOI: 10.3389/fsufs.2020.00013.

McGrenere, Joanna and Wayne Ho (2000). "Affordances: Clarifying and Evolving a Concept". In: *Proceedings of the Graphics Interface 2000 Conference*, pp. 179–186. DOI: 10.20380/GI2000.24.

McKnight, D. Harrison, Michelle Carter, et al. (2011). "Trust in a Specific Technology: An Investigation of Its Components and Measures". In: *ACM*

Transactions on Management Information Systems 2.2, pp. 1–25. DOI: 10.1145/1985347.1985353.

McKnight, D. Harrison, Vivek Choudhury, and Charles Kacmar (2002). "Developing and Validating Trust Measures for e-Commerce: An Integrative Typology". In: *Information Systems Research* 13.3, pp. 334–359. DOI: 10.1287/isre.13.3.334.81.

McKnight, D. Harrison, Larry L. Cummings, and Norman L. Chervany (1998). "Initial Trust Formation in New Organizational Relationships". In: *The Academy of Management Review* 23.3, pp. 473–490. DOI: 10.2307/259290.

Meadows, Donella H. (2008). *Thinking in Systems: A Primer*. White River Junction, VT: Chelsea Green Publishing. ISBN: 9781603580557.

Mertens, Donna M. (2014). *Research and Evaluation in Education and Psychology*. 4th ed. Thousand Oaks, CA: SAGE Publications. ISBN: 9781483322605.

Messick, David M. and Roderick M. Kramer (2001). "Trust as a Form of Shallow Morality". In: *Trust in Society*. Ed. by Karen S. Cook. New York City, NY: Russell Sage Foundation. Chap. 3, pp. 89–118. ISBN: 9780871542489.

Metzger, Miriam J. and Andrew J. Flanagin (2013). "Credibility and Trust of Information in Online Environments: The Use of Cognitive Heuristics". In: *Journal of Pragmatics* 59, pp. 210–220. DOI: https://doi.org/10.1016/j.pragma.2013.07.012.

Meyer, Carsten, Patrick Weigelt, and Holger Kreft (2016). "Multidimensional Biases, Gaps and Uncertainties in Global Plant Occurrence Information". In: *Ecology Letters* 19.8, pp. 992–1006. DOI: 10.1111/ele.12624.

Meyerson, Debra, Karl E. Weick, and Roderick M. Kramer (1996). "Swift Trust and Temporary Groups". In: *Trust in Organizations: Frontiers of Theory and Research*. Ed. by Roderick M. Kramer and Tom Tyler. Thousand Oaks, CA: SAGE Publications, Inc., pp. 166–195. DOI: 10.4135/9781452243610.n9.

Michener, William K. (2000). "Research Design: Translating Ideas to Data". In: *Ecological Data: Design, Management and Processing*. Ed. by William K.

Michener and James W. Brunt. Oxford, UK: Wiley-Blackwell. Chap. 1, pp. 1–24. ISBN: 9780632052318.

Mishler, Elliot G. (1990). "Validation in Inquiry-Guided Research: The Role of Exemplars in Narrative Studies". In: *Harvard Educational Review* 60.4, pp. 415–442. DOI: 10.17763/haer.60.4.n4405243p6635752.

Misztal, Barbara A. (2020). "Trust and the Variety of Its Bases". In: *The Cambridge Handbook of Social Theory*. Ed. by Peter Kivisto. Vol. 2. Cambridge University Press, pp. 333–353. DOI: 10.1017/9781316677452.018.

Møller, Naja Holten et al. (2020). "Who Does the Work of Data?" In: *Interactions* 27.3, pp. 52–55. DOI: 10.1145/3386389.

Möllering, Guido (2001). "The Nature of Trust: From Georg Simmel to a Theory of Expectation, Interpretation and Suspension". In: *Sociology* 35.2, pp. 403–420. DOI: 10.1177/S0038038501000190.

— (2005). *Understanding Trust from the Perspective of Sociological Neoinstitutionalism: The Interplay of Institutions and Agency*. MPIfG Discussion Paper 05/13. Accessed 27/11/20. Cologne. URL: http://hdl.handle.net/10419/19927.

— (2006). "Trust, Institutions, Agency: Towards a Neoinstitutional Theory of Trust". In: *Handbook of Trust Research*. Ed. by Reinhard Bachmann and Akbar Zaheer. Cheltenham, UK: Edward Elgar. Chap. 20, pp. 355–376. ISBN: 9781847202819.

— (2008). *Inviting or Avoiding Deception through Trust? Conceptual Exploration of an Ambivalent Relationship*. Working Paper. Accessed 27/11/21. Max Planck Institute for the Study of Societies, Cologne. URL: https://core.ac.uk/download/pdf/71726373.pdf.

— (2013). "Process Views of Trusting and Crises". In: *Handbook of Advances in Trust Research*. Ed. by Reinhard Bachmann and Akbar Zaheer. Cheltenham, UK: Edward Elgar. Chap. 12, pp. 285–306. ISBN: 9780857931382.

Moses, Jonathon W. and Torbjörn L. Knutsen (2019). *Ways of Knowing: Competing Methodologies in Social and Political Research*. 3rd ed. London: Macmillan Education. ISBN: 9781352005547.

Muir, Bonnie M. and Neville Moray (1996). "Trust in Automation. Part II. Experimental Studies of Trust and Human Intervention in a Process Control Simulation". In: *Ergonomics* 39.3, pp. 429–460. DOI: 10.1080/00140139608964474.

Myskja, Bjørn K. (2008). "The Categorical Imperative and the Ethics of Trust". In: *Ethics and Information Technology* 10, pp. 213–220. DOI: 10.1007/s10676-008-9173-7.

Natural Environment Research Council (2022). *Digitally Enabled Environmental Science: NERC Digital Strategy 2021-2030*. Tech. rep. Accessed 25/06/23. URL: https://www.ukri.org/wp-content/uploads/2022/05/NERC-170522-NERCDigitalStrategy-FINAL-WEB.pdf.

Neff, Gina et al. (2017). "Critique and Contribute: A Practice-Based Framework for Improving Critical Data Studies and Data Science". In: *Big Data* 5.2, pp. 85–97. DOI: 10.1089/big.2016.0050.

Newell, Susan and Jacky Swan (2000). "Trust and Inter-organizational Networking". In: *Human Relations* 53.10, pp. 1287–1328. DOI: 10.1177/a014106.

Newman-Norlund, Sarah E. et al. (2009). "Recipient Design in Tacit Communication". In: *Cognition* 111.1, pp. 46–54. DOI: https://doi.org/10.1016/j.cognition.2008.12.004.

Nickel, Philip J., Maarten Franssen, and Peter Kroes (2010). "Can We Make Sense of the Notion of Trustworthy Technology?" In: *Knowledge, Technology & Policy* 23.3, pp. 429–444. DOI: 10.1007/s12130-010-9124-6.

Nissenbaum, Helen (2001). "Securing Trust Online: Wisdom or Oxymoron?" In: *Boston University Law Review* 81.3. Accessed 26/10/21, pp. 635–664. URL: https://nissenbaum.tech.cornell.edu/papers/securingtrust.pdf.

Noriega, Pablo et al. (2015). "Towards a Framework for Socio-Cognitive Technical Systems". In: *Coordination, Organizations, Institutions, and Norms in Agent*

Systems X. Ed. by Aditya Ghose et al. Cham: Springer International Publishing, pp. 164–181. URL: https://doi.org/10.1007/978-3-319-25420-3_11.

Norman, Donald A. (1988). *The Psychology of Everyday Things*. New York, NY: Basic Books. ISBN: 9780465067091.

— (1999). "Affordance, Conventions, and Design". In: *Interactions* 6.3, pp. 38–43. DOI: 10.1145/301153.301168.

— (2002). *The Design of Everyday Things*. New York, NY: Basic Books. ISBN: 0465067107.

Norman, Donald A. and Stephen W. Draper (1986). *User Centered System Design; New Perspectives on Human-Computer Interaction*. Hillsdale, NJ: L. Erlbaum Associates Inc. ISBN: 0898597811.

Norval, Chris et al. (2022). "Disclosure by Design: Designing Information Disclosures to Support Meaningful Transparency and Accountability". In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '22. Seoul, Republic of Korea, pp. 679–690. DOI: 10.1145/3531146.3533133.

Nowell, Lorelli S. et al. (2017). "Thematic Analysis: Striving to Meet the Trustworthiness Criteria". In: *International Journal of Qualitative Methods* 16.1, pp. 1–13. DOI: 10.1177/1609406917733847.

O'Hara, Kieron (2019). *Data Trusts: Ethics, Architecture and Governance for Trustworthy Data Stewardship*. Project Report. DOI: 10.5258/SOTON/WSI-WP001.

O'Neill, Onora (2002). *A Question of Trust: The BBC Reith Lectures*. Cambridge, UK: Cambridge University Press. ISBN: 978-0521529969.

— (2018). "Linking Trust to Trustworthiness". In: *International Journal of Philosophical Studies* 26.2, pp. 293–300. DOI: 10.1080/09672559.2018.1454637.

O'Day, Vicki L. and Bonnie A. Nardi (2003). "An Ecological Perspective on Digital Libraries". In: *Digital Library Use: Social Practice in Design and Evaluation*. Ed. by Ann Peterson Bishop, Nancy A. Van House, and Barbara P. Buttenfield. Cambridge, MA: The MIT Press. Chap. 4, pp. 65–82. ISBN: 0262025442.

O'Neill, Onora (2020). "Questioning Trust". In: *The Routledge Handbook of Trust and Philosophy*. Ed. by Judith Simon. New York, NY: Routledge. Chap. 1, pp. 17–27. ISBN: 9781315542294.

Obar, Jonathan A. (2020). "Sunlight Alone is not A Disinfectant: Consent and the Futility of Opening Big Data Black Boxes (Without Assistance)". In: *Big Data & Society* 7.1, p. 2053951720935615. DOI: 10.1177/2053951720935615.

Oliver, Martin (2005). "The Problem with Affordance". In: *E-Learning and Digital Media* 2.4, pp. 402–413. DOI: 10.2304/elea.2005.2.4.402.

Olson, Gary M. and Judith S. Olson (2000). "Distance Matters". In: *Human-Computer Interaction* 15.2, pp. 139–178. DOI: 10.1207/S15327051HCI1523_4.

Olson, Gary M., Stephanie Teasley, et al. (2002). "Collaboratories to Support Distributed Science: The Example of International HIV/AIDS Research". In: *Proceedings of the 2002 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on Enablement through Technology*. SAICSIT '02. Port Elizabeth, South Africa: South African Institute for Computer Scientists and Information Technologists, pp. 44–51. URL: https://dl.acm.org/doi/10.5555/581506.581512.

Oreskes, Naomi (2019). *Why Trust Science?* Princeton, NJ: Princeton University Press. ISBN: 9780691179001. URL: http://www.jstor.org/stable/j.ctvfjczxx.

Ostern, Nadine (2018). "Do you Trust a Trust-free Transaction? Toward a Trust Framework Model for Blockchain Technology". In: *Thirty Ninth International Conference on Information Systems*. Accessed 19/03/20. URL: https://aisel.aisnet.org/icis2018/crypto/Presentations/3/.

Ostern, Nadine Kathrin (2020). "Blockchain in the IS Research Discipline: A Discussion of Terminology and Concepts". In: *Electronic Markets* 30.2, pp. 195–210. URL: https://doi.org/10.1007/s12525-019-00387-2.

Pasquetto, Irene V., Christine L. Borgman, and Morgan F. Wofford (2019). "Uses and Reuses of Scientific Data: The Data Creators' Advantage". In: *Harvard Data Science Review* 1.2 (2). DOI: 10.1162/99608f92.fc14bf2d.

Passi, Samir and Steven J. Jackson (2018). "Trust in Data Science: Collaboration, Translation, and Accountability in Corporate Data Science Projects". In: *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW, pp. 1–28. DOI: 10.1145/3274405.

Patrick, Andrew S., Pamela Briggs, and Stephen Marsh (2005). "Designing Systems That People Will Trust". In: *Security and Usability: Designing Secure Systems that People Can Use*. Ed. by Lorrie Faith Cranor and Simson Garfinkel. Sebastopol, CA: O'Reilly. Chap. 5, pp. 75–100. ISBN: 9780596008277.

Patton, Michael Quinn (2002). *Qualitative Research & Evaluation Methods*. 3rd ed. Thousand Oaks, CA: Sage Publications. ISBN: 0761919716.

Penner, David E. and David Klahr (1996). "When to Trust the Data: Further Investigations of System Error in a Scientific Reasoning Task". In: *Memory and Cognition* 24.5, pp. 655–668. ISSN: 0090502X. DOI: 10.3758/BF03201090.

Pescott, Oliver, Thomas Humphrey, and Kevin Walker (July 2018). "A Short Guide to Using British and Irish Plant Occurrence Data for Research". In: Accessed 20/05/20. URL: https://nora.nerc.ac.uk/id/eprint/520609/1/N520609CR.pdf.

Pettit, Philip (1995). "The Cunning of Trust". In: *Philosophy & Public Affairs* 24.3, pp. 202–225. DOI: 10.1111/j.1088-4963.1995.tb00029.x.

Petty, Nicola J., Oliver P. Thomson, and Graham Stew (2012). "Ready for a Paradigm Shift? Part 2: Introducing Qualitative Research Methodologies and Methods". In: *Manual Therapy* 17.5, pp. 378–384. DOI: 10.1016/j.math.2012.03.004.

Pieters, Wolter (2011). "Explanation and Trust: What to Tell the User in Security and AI?" In: *Ethics and Inf. Technol.* 13.1, pp. 53–64. DOI: 10.1007/s10676-010-9253-3.

Pink, Sarah, Debora Lanzeni, and Heather Horst (2018). "Data Anxieties: Finding Trust in Everyday Digital Mess". In: *Big Data & Society* 5.1. DOI: 10.1177/2053951718756685.

Piorkowski, David et al. (2021). "How AI Developers Overcome Communication Challenges in a Multidisciplinary Team: A Case Study". In: *Proceedings of the ACM on Human-Computer Interaction* 5. DOI: 10.1145/3449205.

Pirson, Michael and Deepak Malhotra (2010). "Foundations of Organizational Trust: What Matters to Different Stakeholders?" In: *Organization Science* 22.4, pp. 1087–1104. DOI: 10.1287/orsc.1100.0581.

Poursabzi-Sangdeh, Forough et al. (2021). "Manipulating and Measuring Model Interpretability". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Yokohama, Japan. DOI: 10.1145/3411764.3445315.

Preece, Alun D. et al. (2018). "Stakeholders in Explainable AI". In: *CoRR* abs/1810.00184. arXiv: 1810.00184.

Prieto, Adolfo G. (2009). "From Conceptual to Perceptual Reality: Trust in Digital Repositories". In: *Library Review* 58.8, pp. 593–606. DOI: 10.1108/00242530910987082.

Priya, Arya (2021). "Case Study Methodology of Qualitative Research: Key Attributes and Navigating the Conundrums in Its Application". In: *Sociological Bulletin* 70.1, pp. 94–110. DOI: 10.1177/0038022920970318.

Raji, Inioluwa Deborah et al. (2020). "Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing". In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. FAT* '20. Barcelona, Spain, pp. 33–44. DOI: 10.1145/3351095.3372873.

Ramírez-i-Ollé, Meritxell (2018). "'Civil Skepticism' and the Social Construction of Knowledge: A Case in Dendroclimatology". In: *Social Studies of Science* 48.6, pp. 821–845. DOI: 10.1177/0306312718763119.

Rashid, Yasir et al. (2019). "Case Study Method: A Step-by-Step Guide for Business Researchers". In: *International Journal of Qualitative Methods* 18, pp. 1–13. DOI: 10.1177/1609406919862424.

Reis, Stefan et al. (2015). "Integrating Modelling and Smart Sensors for Environmental and Human Health". In: *Environmental Modelling & Software* 74, pp. 238–246. DOI: 10.1016/j.envsoft.2015.06.003.

Richards, John T. et al. (2021). "A Human-Centered Methodology for Creating AI FactSheets". In: *IEEE Database Engineering Bulletin* 44 (4). Accessed 05/01/22, pp. 47–58. URL: https://krvarshney.github.io/pubs/RichardsPHHMV_btcde2021.pdf.

Richards, Lyn (2005). *Handling Qualitative Data: A Practical Guide*. 3rd ed. London: SAGE Publications. ISBN: 9780761942580.

Rieder, Gernot and Judith Simon (2016). "Datatrust: Or, the Political Quest for Numerical Evidence and the Epistemologies of Big Data". In: *Big Data & Society* 3.1, pp. 1–6. DOI: 10.1177/2053951716649398.

Riegelsberger, Jens (2002). "The Effect of Facial Cues on Trust in E-commerce Systems". In: *Proceedings of the 16th British HCI Conference, Volume II*. Accessed 06/01/22. London, UK. URL: http://www0.cs.ucl.ac.uk/staff/j.riegelsberger/facialcues.pdf.

Riegelsberger, Jens and M. Angela Sasse (2010). "Ignore These At Your Peril: Ten Principles for Trust Design". In: *3rd International Conference on Trust and Trustworthy Computing*. Accessed 08/12/21. URL: https://discovery.ucl.ac.uk/id/eprint/20297/.

Riegelsberger, Jens, M. Angela Sasse, and John D. McCarthy (2003a). "Shiny Happy People Building Trust? Photos on e-Commerce Websites and Consumer Trust". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '03. Ft. Lauderdale, Florida, USA: Association for Computing Machinery, pp. 121–128. ISBN: 1581136307. DOI: 10.1145/642611.642634.

— (2005). "The Mechanics of Trust: A Framework for Research and Design". In: *International Journal of Human-Computer Studies* 62.3, pp. 381–422. DOI: 10.1016/j.ijhcs.2005.01.001.

— (2003b). "The Researcher's Dilemma: Evaluating Trust in Computer-Mediated Communication". In: *International Journal of Human-Computer Studies* 58.6, pp. 759–781. DOI: https://doi.org/10.1016/S1071-5819(03)00042-9.

Ritchie, Jane and Jane Lewis (2003). *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. London: Sage Publications. ISBN: 0761971092.

Robbins, Blaine G. (2016). "From the General to the Specific: How Social Trust Motivates Relational Trust". In: *Social Science Research* 55, pp. 16–30. DOI: 10.1016/j.ssresearch.2015.09.004.

Rodden, Tom and Gordon Blair (1991). "CSCW and Distributed Systems: The Problem of Control". In: *Proceedings of the Second European Conference on Computer-Supported Cooperative Work ECSCW '91*. Ed. by Liam Bannon, Mike Robinson, and Kjeld Schmidt. Dordrecht, NL: Springer Netherlands, pp. 49–64. DOI: 10.1007/978-94-011-3506-1_4.

Rolin, Kristina (2020). "Trust in Science". In: *The Routledge Handbook of Trust and Philosophy*. Ed. by Judith Simon. New York, NY: Routledge. Chap. 27, pp. 354–366. ISBN: 9781315542294.

Ross, Seamus and Andrew Mchugh (2006). "The Role of Evidence in Establishing Trust in Repositories". In: *D-Lib Magazine* 12 (7/8). DOI: 10.1045/july2006-ross.

Rothenberg, Jeff (2000). "Presenting Authentic Digital Information". In: *Authenticity in a Digital Environment*. Ed. by Charles T. Cullen et al. Council on Library and Information Resources. Chap. 5, pp. 51–68. ISBN: 1887334777.

Rousseau, Denise M. et al. (1998). "Not So Different After All: A Cross-Discipline View of Trust". In: *Academy of Management Review* 23.3, pp. 393–404. DOI: 10.5465/amr.1998.926617.

Roy, Kevin et al. (2015). "Sampling Richness and Qualitative Integrity: Challenges for Research With Families". In: *Journal of Marriage and Family* 77.1, pp. 243–260. DOI: 10.1111/jomf.12147.

Rubin, Herbert J. and Irene S. Rubin (2005). *Qualitative Interviewing: The Art of Hearing Data*. 2nd ed. Thousand Oaks, CA: SAGE Publications. DOI: 10.4135/9781452226651.

Sabater, Jordi and Carles Sierra (2005). "Review on Computational Trust and Reputation Models". In: *Artificial Intelligence Review* 24.1, pp. 33–60. ISSN: 1573-7462. DOI: 10.1007/s10462-004-0041-5.

Saldaña, Johnny (2013). *The Coding Manual for Qualitative Researchers*. 2nd ed. London, UK: SAGE. ISBN: 9781847875495.

Sandelowski, Margarete (1996). "One is the Liveliest Number: The Case Orientation of Qualitative Research". In: *Research in Nursing & Health* 19.6, pp. 525–529. URL: https://doi.org/10.1002/(SICI)1098-240X(199612)19:6%3C525::AID-NUR8%3E3.0.CO;2-Q.

Sandelowski, Margarete and Julie Barroso (2002). "Reading Qualitative Studies". In: *International Journal of Qualitative Methods* 1.1, pp. 74–108. DOI: 10.1177/160940690200100107.

Sas, Corina and Irni Eliana Khairuddin (2015). "Exploring Trust in Bitcoin Technology: A Framework for HCI Research". In: *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*. OzCHI '15. Parkville, VIC, Australia: Association for Computing Machinery, pp. 338–342. ISBN: 9781450336734. DOI: 10.1145/2838739.2838821.

— (2017). "Design for Trust: An Exploration of the Challenges and Opportunities of Bitcoin Users". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, pp. 6499–6510. ISBN: 9781450346559. DOI: 10.1145/3025453.3025886.

Saunders, Mark N. K., Philip Lewis, and Adrian Thornhill (2000). *Research Methods for Business Students*. 2nd ed. Harlow, Essex: Pearson Education. ISBN: 0273639773.

Scheman, Naomi (2020). "Trust and Trustworthiness". In: *The Routledge Handbook of Trust and Philosophy*. Ed. by Judith Simon. New York, NY: Routledge. Chap. 2, pp. 28–41. ISBN: 9781315542294.

Schiff, Lisa R., Nancy A. Van House, and Mark H. Butler (1997). "Understanding Complex Information Environments: A Social Analysis of Watershed Planning". In: *Proceedings of the Second ACM International Conference on Digital Libraries*. DL '97. Philadelphia, Pennsylvania, USA: Association for Computing Machinery, pp. 161–168. ISBN: 0897918681. DOI: 10.1145/263690.263809.

Schilke, Oliver and Karen S. Cook (2013). "A Cross-level Process Theory of Trust Development in Interorganizational Relationships". In: *Strategic Organization* 11.3, pp. 281–303. DOI: 10.1177/1476127012472096.

Schmidt, Kjeld (2000). "The critical role of workplace studies in CSCW". In: *Workplace Studies: Recovering Work Practice and Informing System Design*. Ed. by Paul Luff, Jon Hindmarsh, and Christian Heath. Cambridge, UK: Cambridge University Press. Chap. 6, pp. 141–149. ISBN: 0521598214.

Schmidt, Kjeld and Liam Bannon (1992). "Taking CSCW Seriously: Supporting Articulation Work". In: *Computer Supported Cooperative Work* 1.1, pp. 7–40. DOI: 10.1007/BF00752449.

Schoorman, F. David, Roger C. Mayer, and James H. Davis (2007). "An Integrative Model of Organizational Trust: Past, Present, and Future". In: *Academy of Management Review* 32.2, pp. 344–354. DOI: 10.5465/amr.2007.24348410.

Schutt, Russell K. (2012). *Investigating the Social World: The Process and Practice of Research*. 7th ed. Thousand Oaks, CA: SAGE Publications. ISBN: 9781412999809.

Sellen, Abigail J. and Richard H.R. Harper (2001). *The Myth of the Paperless Office*. Cambridge, MA: MIT Press. ISBN: 9780262194648.

Sengers, Phoebe and Bill Gaver (2006). "Staying Open to Interpretation: Engaging Multiple Meanings in Design and Evaluation". In: *Proceedings of the 6th Conference on Designing Interactive Systems*. Association for Computing Machinery, pp. 99–108. DOI: 10.1145/1142405.1142422.

Sexton, Anna et al. (2005). "A Balance of Trust in the Use of Government Administrative Data". In: *Archival Science* 17.4, pp. 305–330. ISSN: 1573-7519. DOI: 10.1007/s10502-017-9281-4.

Shannon, Claude E. (1948). "A Mathematical Theory of Communication". In: *The Bell System Technical Journal* 27.3, pp. 379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

Shapiro, Susan P. (1987). "The Social Control of Impersonal Trust". In: *American Journal of Sociology* 93.3, pp. 623–658. URL: https://www.jstor.org/stable/2780293.

Sharp, Emily A. et al. (2013). "Trust and Trustworthiness: Conceptual Distinctions and their Implications for Natural Resources Management". In: *Journal of Environmental Planning and Management* 56.8, pp. 1246–1265. DOI: 10.1080/09640568.2012.717052.

Shin, Donghee, Bu Zhong, and Frank A. Biocca (2020). "Beyond User Experience: What Constitutes Algorithmic Experiences?" In: *International Journal of Information Management* 52. DOI: https://doi.org/10.1016/j.ijinfomgt.2019.102061.

Simmel, Georg (1950). *The Sociology of Georg Simmel*. New York, NY: Free Press. ISBN: 9780029289204.

Simon, Judith (2020). "Introduction". In: *The Routledge Handbook of Trust and Philosophy*. Ed. by Judith Simon. New York, NY: Routledge, pp. 1–13. ISBN: 9781315542294.

Simpson, Thomas W. (2012). "What Is Trust?" In: *Pacific Philosophical Quarterly* 93.4, pp. 550–569. DOI: 10.1111/j.1468-0114.2012.01438.x.

— (2014). "Computing and the Search for Trust". In: *Trust, Computing, and Society*. Ed. by Richard H. Harper. Cambridge: Cambridge University Press. Chap. 5, pp. 95–119. DOI: 10.1017/CBO9781139828567.009.

Sitkin, Sim B. and Nancy L. Roth (1993). "Explaining the Limited Effectiveness of Legalistic "Remedies" for Trust/Distrust". In: *Organization Science* 4.3, pp. 367–392. DOI: 10.1287/orsc.4.3.367.

Smith, Abby (2000). "Authenticity in Perspective". In: *Authenticity in a Digital Environment*. Ed. by Charles T. Cullen et al. Council on Library and Information Resources. Chap. 6, pp. 69–75. ISBN: 1887334777.

Sokol, Kacper and Peter Flach (2020). "Explainability Fact Sheets: A Framework for Systematic Assessment of Explainable Approaches". In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. FAT* '20. Barcelona, Spain, pp. 56–67. DOI: 10.1145/3351095.3372870.

Söllner, Matthias et al. (Dec. 2012). "Understanding the Formation of Trust in IT Artifacts". In: vol. 2. Accessed 15/11/21. Orlando, FL, pp. 1–18. URL: https://aisel.aisnet.org/icis2012/proceedings/HumanBehavior/11/.

Sommerville, Ian et al. (2006). "Dependability and Trust in Organisational and Domestic Computer Systems". In: *Trust in Technology: A Socio-Technical Perspective*. Ed. by Karen Clarke et al. Dordrecht: Springer Netherlands, pp. 169–193. DOI: 10.1007/1-4020-4258-2_8.

Spiekermann, Raphael et al. (2019). "Implementations of Fine-Grained Automated Data Provenance to Support Transparent Environmental Modelling". In: *Environmental Modelling & Software* 118, pp. 134–145. DOI: 10.1016/j.envsoft.2019.04.009.

Stake, Robert. E (1995). *The Art of Case Study Research*. Thousand Oaks, CA: SAGE Publications. ISBN: 9780803957671.

— (2005). *Multiple Case Study Analysis*. New York, NY: Guilford Press. ISBN: 9781593852481.

Stall, Shelley et al. (2019). "Make Scientific Data FAIR". In: *Nature* 570, pp. 27–29. DOI: 10.1038/d41586-019-01720-7.

Star, Susan Leigh, Geoffrey C. Bowker, and Laura J. Neumann (2003). "Transparency beyond the Individual Level of Scale: Convergence between Information Artifacts and Communities of Practice". In: *Digital Library Use: Social Practice in Design and Evaluation*. Ed. by Ann Peterson Bishop, Nancy A. Van House, and Barbara P. Buttenfield. Cambridge, MA: The MIT Press. Chap. 10, pp. 241–269. ISBN: 0262025442.

Steedman, Robin, Helen Kennedy, and Rhianne Jones (2020). "Complex Ecologies of Trust in Data Practices and Data-driven Systems". In: *Information, Communication & Society* 23.6, pp. 817–832. DOI: 10.1080/1369118X.2020.1748090.

Still, Jeremiah D. and Veronica J. Dark (2013). "Cognitively Describing and Designing Affordances". In: *Design Studies* 34.3, pp. 285–301. DOI: 10.1016/j.destud.2012.11.005.

Sturdee, Miriam et al. (2021). "A Visual Exploration of Cybersecurity Concepts". In: *C&C '21: Creativity and Cognition*. ACM. DOI: 10.1145/3450741.3465252.

Suchman, Lucy A. (1987). *Plans and Situated Actions: The Problem of Human-Machine Communication*. USA: Cambridge University Press. ISBN: 0521331374.

— (2002). "Practice-Based Design of Information Systems: Notes from the Hyperdeveloped World". In: *The Information Society* 18.2, pp. 139–144. DOI: 10.1080/01972240290075066.

Sundqvist, Anneli (2011). "Documentation Practices and Recordkeeping: A Matter of Trust or Distrust?" In: *Archival Science* 11.4, pp. 277–291. DOI: 10.1007/s10502-011-9160-3.

Tabuenca, Bernardo et al. (2021). "Affordances and Core Functions of Smart Learning Environments: A Systematic Literature Review". In: *IEEE Transactions on Learning Technologies* 14.2, pp. 129–145. DOI: 10.1109/TLT.2021.3067946.

Talboom, Sarah and Jo Pierson (2013). "Understanding Trust within Online Discussion Boards: Trust Formation in the Absence of Reputation Systems".

In: *Trust Management VII*. Ed. by Carmen Fernández-Gago et al. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 83–99. DOI: 10.1007/978-3-642-38323-6_6.

Thornton, Lauren, Bran Knowles, and Gordon Blair (2021). "Fifty Shades of Grey: In Praise of a Nuanced Approach Towards Trustworthy Design". In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '21. Virtual Event, Canada: Association for Computing Machinery, pp. 64–76. ISBN: 9781450383097. DOI: 10.1145/3442188.3445871.

— (2022). "The Alchemy of Trust: The Creative Act of Designing Trustworthy Socio-Technical Systems". In: *2022 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '22. Seoul, Republic of Korea: Association for Computing Machinery, pp. 1387–1398. ISBN: 9781450393522. DOI: 10.1145/3531146.3533196.

Thornton, Lauren, Victoria Neumann, et al. (2019). "Trusted Brokers?: Identifying the Challenges Facing Data Centres". In: Data for Policy. DOI: 10.5281/zenodo.2798468.

Tillas, Alexandros et al. (2017). "Can Affordances Explain Behavior?" In: *Review of Philosophy and Psychology* 8.2, pp. 295–315. DOI: 10.1007/s13164-016-0310-7.

Torenvliet, Gerard (2003). "We Can't Afford It!: The Devaluation of a Usability Term". In: *Interactions* 10.4, pp. 12–17. DOI: 10.1145/838830.838857.

Tracy, Sarah J. (2010). "Qualitative Quality: Eight "Big-Tent" Criteria for Excellent Qualitative Research". In: *Qualitative Inquiry* 16.10, pp. 837–851. DOI: 10.1177/1077800410383121.

Turilli, Matteo, Antonino Vaccaro, and Mariarosaria Taddeo (2010). "The Case of Online Trust". In: *Knowledge, Technology & Policy* 23, pp. 333–345. DOI: 10.1007/s12130-010-9117-5.

Underwood, Sarah (2016). "Blockchain Beyond Bitcoin". In: *Communications of the ACM* 59.11, pp. 15–17. DOI: 10.1145/2994581.

Uslaner, Eric M. (2000). "Producing and Consuming Trust". In: *Political Science Quarterly* 115.4, pp. 569–590. URL: http://www.jstor.org/stable/2657610.

— (2018). "The Study of Trust". In: *The Oxford Handbook of Social and Political Trust*. Ed. by Eric M. Uslaner. Oxford, UK: Oxford University Press. Chap. 1, pp. 3–14. DOI: 10.1093/oxfordhb/9780190274801.013.39.

Uther, Maria and Adrian P. Banks (2016). "The Influence of Affordances on User Preferences for Multimedia Language Learning Applications". In: *Behaviour & Information Technology* 35.4, pp. 277–289. DOI: 10.1080/0144929X.2016.1151077.

Vakarelov, Orlin (2012). "The Information Medium". In: *Philosophy & Technology* 25.1, pp. 47–65. DOI: 10.1007/s13347-011-0016-9.

Van House, Nancy A. (2002). "Digital Libraries and Practices of Trust: Networked Biodiversity Information". In: *Social Epistemology* 16.1, pp. 99–114. DOI: 10.1080/02691720210132833.

— (2003). "Digital Libraries and Collective Knowledge Construction". In: *Digital Library Use: Social Practice in Design and Evaluation*. Ed. by Ann Peterson Bishop, Nancy A. Van House, and Barbara P. Buttenfield. Cambridge, MA: The MIT Press. Chap. 11, pp. 271–295. ISBN: 0262025442.

Van House, Nancy A., Ann Peterson Bishop, and Barbara P. Buttenfield (2003). "Digital Libraries as Sociotechnical Systems". In: *Digital Library Use: Social Practice in Design and Evaluation*. Ed. by Ann Peterson Bishop, Nancy A. Van House, and Barbara P. Buttenfield. Cambridge, MA: The MIT Press. Chap. 1, pp. 1–24. ISBN: 0262025442.

Van House, Nancy A., Mark H. Butler, Virginia E. Ogle, et al. (1996). "User-Centered Iterative Design for Digital Libraries: The Cypress Experience". In: *D Lib Mag.* 2. Accessed 28/11/21. URL: http://www.dlib.org/dlib/february96/02vanhouse.html.

Van House, Nancy A., Mark H. Butler, and Lisa R. Schiff (1998). "Cooperative Knowledge Work and Practices of Trust: Sharing Environmental Planning Data

Sets". In: *CSCW '98, Proceedings of the ACM 1998 Conference on Computer Supported Cooperative Work*. Ed. by Steven E. Poltrock and Jonathan Grudin. ACM, pp. 335–343. DOI: 10.1145/289444.289508.

Vasileiou, Konstantina et al. (2018). "Characterising and Justifying Sample Size Sufficiency in Interview-based Studies: Systematic Analysis of Qualitative Health Research over a 15-year period". In: *BMC Medical Research Methodology* 18.148, pp. 1–18. DOI: 10.1186/s12874-018-0594-7.

Vaughan, Jennifer Wortman and Hanna Wallach (2021). "A Human-Centered Agenda for Intelligible Machine Learning". In: *Machines We Trust: Perspectives on Dependable AI*. Ed. by Marcello Pelillo and Teresa Scantamburlo. Cambridge, MA: The MIT Press. Chap. 8, pp. 123–138. ISBN: 9780262542098.

Völter, Fabiane, Nils Urbach, and Julian Padget (2023). "Trusting the Trust Machine: Evaluating Trust Signals of Blockchain Applications". In: *International Journal of Information Management* 68, pp. 1–13. DOI: https://doi.org/10.1016/j.ijinfomgt.2021.102429.

Voss, Alex et al. (2009). "Design as and for Collaboration: Making Sense of and Supporting Practical Action". In: *Configuring User-Designer Relations: Interdisciplinary Perspectives*. Ed. by Alex Voss et al. London: Springer London. Chap. 3, pp. 31–58. ISBN: 978-1-84628-925-5. DOI: 10.1007/978-1-84628-925-5_3.

Vyas, Dhaval, Cristina M. Chisalita, and Alan Dix (2017). "Organizational Affordances: A Structuration Theory Approach to Affordances". In: *Interacting with Computers* 29.2, pp. 117–131. DOI: 10.1093/iwc/iww008.

Vyas, Dhaval, Cristina M. Chisalita, and Gerrit C. van der Veer (2006). "Affordance in Interaction". In: *Proceedings of the 13th Eurpoean Conference on Cognitive Ergonomics: Trust and Control in Complex Socio-Technical Systems*. Zurich, Switzerland: Association for Computing Machinery, pp. 92–99. DOI: 10.1145/1274892.1274907.

Wadlow, Thomas (2014). "Who Must You Trust?" In: *Communications of the ACM* 57.7, pp. 42–49. DOI: 10.1145/2619238.

Wallis, Jillian C., Christine L. Borgman, et al. (2007). "Know Thy Sensor: Trust, Data Quality, and Data Integrity in Scientific Digital Libraries". In: *Research and Advanced Technology for Digital Libraries*. Vol. 4675, pp. 380–391. DOI: 10.1007/978-3-540-74851-9_32.

Wallis, Jillian C., Elizabeth Rolando, and Christine L. Borgman (2013). "If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology". In: *PLOS ONE* 8.7, pp. 1–17. DOI: 10.1371/journal.pone.0067332.

Weltevrede, Esther and Erik Borra (2016). "Platform Affordances and Data Practices: The Value of Dispute on Wikipedia". In: *Big Data & Society* 3.1, pp. 1–16. DOI: 10.1177/2053951716653418.

Wilson, William J. and Anmol Chaddha (2009). "The Role of Theory in Ethnographic Research". In: *Ethnography* 10.4, pp. 549–564. DOI: 10.1177/1466138109347009.

Winters, Niall and Yishay Mor (2009). "Dealing with Abstraction: Case Study Generalisation as a Method for Eliciting Design Patterns". In: *Computers in Human Behavior* 25.5. Including the Special Issue: Design Patterns for Augmenting E-Learning Experiences, pp. 1079–1088. DOI: 10.1016/j.chb.2009.01.007.

Yakel, Elizabeth et al. (2013). "Trust in Digital Repositories". In: *International Journal of Digital Curation* 8 (1), pp. 143–156. URL: https://ijdc.net/index.php/ijdc/article/view/8.1.143.

Yin, Robert K. (2017). *Case Study Research and Applications: Design and Methods*. 6th ed. Thousand Oaks, CA: SAGE Publications. ISBN: 9781506336183.

Yoon, Ayoung (2014). "End Users' Trust in Data Repositories: Definition and Influences on Trust Development". In: *Archival Science* 14, pp. 17–34. DOI: 10.1007/s10502-013-9207-8.

— (2017). "Data Reusers' Trust Development". In: *Journal of the Association for Information Science and Technology* 68 (4), pp. 946–956. DOI: 10.1002/asi.23730.

Zaheer, Akbar, Bill McEvily, and Vincenzo Perrone (1998). "Does Trust Matter? Exploring the Effects of Interorganizational and Interpersonal Trust on Performance". In: *Organization Science* 9.2, pp. 141–159. DOI: 10.1287/orsc.9.2.141.

Zhao, Jun et al. (2011). "eScience". In: *Handbook of Semantic Web Technologies*. Ed. by John Domingue, Dieter Fensel, and James A. Hendler. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 701–736. ISBN: 978-3-540-92913-0. DOI: 10.1007/978-3-540-92913-0_17.

Zimmerman, Ann and Bonnie A. Nardi (2006). "Whither or Whether HCI: Requirements Analysis for Multi-Sited, Multi-User Cyberinfrastructures". In: *CHI '06 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '06. Montréal, Québec, Canada: Association for Computing Machinery, pp. 1601–1606. DOI: 10.1145/1125451.1125743.

Zimmerman, Ann S. (2008a). "New Knowledge from Old Data: The Role of Standards in the Sharing and Reuse of Ecological Data". In: *Science, Technology, & Human Values* 33.5, pp. 631–652. DOI: 10.1177/0162243907306704.

— (2008b). "Not by Metadata Alone: The Use of Diverse Forms of Knowledge to Locate Data for Reuse". In: *International Journal on Digital Libraries* 7.1, pp. 5–16. DOI: 10.1007/s00799-007-0015-8.

Zimmerman, Ann S., Nathan Bos, et al. (2009). "The Promise of Data in e-Research: Many Challenges, Multiple Solutions, Diverse Outcomes". In: *E-Research: Transformation in Scholarly Practice*. Ed. by Nicholas W. Jankowski. New York, NY: Routledge. Chap. 12, pp. 222–239. ISBN: 9780203875049.

Zucker, Lynne G. (1986). "Production of Trust: Institutional Sources of Economic Structure, 1840–1920". In: *Research in Organizational Behavior* 8. Accessed

19/10/2021, pp. 53–111. URL: https://digitalassets.lib.berkeley.edu/irle/ucb/text/irla0277.pdf.

Zuiderwijk, Anneke (2017). "Analysing Open Data in Virtual Research Environments: New Collaboration Opportunities to Improve Policy Making". In: *International Journal of Electronic Government Research* 13.4, pp. 76–92. DOI: 10.4018/IJEGR.2017100105.