

# Fifty Shades of Grey: In Praise of a Nuanced Approach Towards Trustworthy Design

Lauren Thornton  
l.thornton2@lancaster.ac.uk  
Lancaster University  
Lancaster, UK

Bran Knowles  
b.h.knowles1@lancaster.ac.uk  
Lancaster University  
Lancaster, UK

Gordon Blair  
g.blair@lancaster.ac.uk  
Lancaster University  
Lancaster, UK

## ABSTRACT

Environmental data science is uniquely placed to respond to essentially complex and fantastically worthy challenges related to arresting planetary destruction. Trust is needed for facilitating collaboration between scientists who may share datasets and algorithms, and for crafting appropriate science-based policies. Achieving this trust is particularly challenging because of the numerous complexities, multi-scale variables, interdependencies and multi-level uncertainties inherent in environmental data science. Virtual Labs—easily accessible online environments provisioning access to datasets, analysis and visualisations—are socio-technical systems which, if carefully designed, might address these challenges and promote trust in a variety of ways. In addition to various system properties that can be utilised in support of effective collaboration, certain features which are commonly seen to benefit trust—transparency and provenance in particular—appear applicable to promoting trust in and through Virtual Labs. Attempting to realise these features in their design reveals, however, that their implementation is more nuanced and complex than it would appear. Using the lens of affordances, we argue for the need to carefully articulate these features, with consideration of multiple stakeholder needs on balance, so that these Virtual Labs do in fact promote trust. We argue that these features not be conceived as widgets that can be imported into a given context to promote trust; rather, whether they promote trust is a function of how systematically designers consider various (potentially conflicting) stakeholder trust needs.

## CCS CONCEPTS

• **Human-centered computing** → **Computer supported cooperative work; HCI theory, concepts and models.**

## KEYWORDS

Trust, Affordances, Environmental Data Science, Virtual Research Environments

### ACM Reference Format:

Lauren Thornton, Bran Knowles, and Gordon Blair. 2021. Fifty Shades of Grey: In Praise of a Nuanced Approach Towards Trustworthy Design. In *Conference on Fairness, Accountability, and Transparency (FAcT '21)*, March

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

FAcT '21, March 3–10, 2021, Virtual Event, Canada

© 2021 Association for Computing Machinery.

ACM ISBN 978-1-4503-8309-7/21/03...\$15.00

<https://doi.org/10.1145/3442188.3445871>

3–10, 2021, Virtual Event, Canada. ACM, New York, NY, USA, 13 pages.  
<https://doi.org/10.1145/3442188.3445871>

## 1 INTRODUCTION

It is well known the numerous benefits of trust, for example in facilitating and sustaining collaborative relationships and enabling decision making [82, 102, 107]. More fundamentally, however, we can understand trust as a clever mechanism, the function of which being to reduce complexity in order to resolve paralysing uncertainty so that one is able to act—or in Möllering’s [111, p. 403] own words: “bracketing the unknowable, thus making interpretative knowledge momentarily certain”. Unknowing is intrinsic to science: we look to evidence to support (or really, the absence of evidence to refute) hypotheses so that we can make confident claims about our world. Confidence in scientific claims is perhaps uniquely challenging in the area of environmental science, where competing values, priorities, and politics are at play [97, 117, 119]. Being able to make policy decisions that will help arrest planetary destruction requires trust 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 [12, 119, 141]. The collaborative nature of science and particularly environmental science also means that scientists themselves need to have a solid basis for trusting their collaborators, the processes and data they use, and their findings, in order to make collective progress in the field.

The turn towards open data and the development of new collaborative environments are indicative of the need to foster trust in environmental data science and its underlying methods, assumptions and datasets. The authors of this paper are involved in the development of an easily accessible online environment provisioning access to datasets, analysis and visualisations. In the early stages of designing this Virtual Lab, it was presumed that well-known trust-supporting properties would need to be incorporated to achieve intended trust [71]:

- (1) *Transparency*: providing collaborators access to relevant information to determine trustworthiness of the items in the repository;
- (2) *Provenance*: preserving workflow through notebook technologies integrated with version control (possibly incorporating blockchain);
- (3) *Security*: balancing openness with some assurance of privacy for collaborators.

It is possible that these are essential high-level design goals in promoting trust *in some form*. But in attempting to envisage how these might be implemented, and interrogating the purpose they serve with respect to specific users of the system in specific

contexts, we came to understand that these properties which are so commonly seen to benefit trust are not as unproblematically associated with trust as the literature would generally have us believe.

This paper offers reflections, grounded in an extensive trust literature and informed by conversations and interviews with relevant stakeholders of this Virtual Lab, on the challenges of simply importing ‘trust solutions’ from other domains. There is neither a singular solution, one-size-fits-all approach, nor a magic formula to designing for trust—this is known already, if not necessarily always appreciated [2, 39, 102]. Where we go further is in problematising two of the most common ingredients used in designing for trust, namely transparency and provenance. Our aim is to inform a more *considered discourse* on what we are calling ‘trust affordances’, i.e. designed aspects of systems that promote trust. We hope to instill a new sensibility about these affordances as being not only context-dependent, but also fluid and multiply-articulated within socio-technical systems. This is quite distinct from the predominant (though implicit) view of such affordances as simple widgets that can be taken off the shelf and incorporated into a system. By exploring specifics of how each affordance contributes to trust, for whom, and then what it would need to look like to promote trust for this particular stakeholder in context, we show that trust affordances would need to vary in their final form even within the same system - in other words, rather than being binary or black and white, there are *many shades of grey*. Given the collaborative nature of environmental data science, we follow on from this with a brief consideration of additional affordances related to the socio-technical aspects of Virtual Labs. We conclude with some thoughts about the applicability of these insights to the wider literature on trust, where transparency and provenance in particular are seen as almost unquestionably beneficial to trust.

## 2 BACKGROUND

### 2.1 Environmental Data Science

Environmental data science applies the tools and techniques of data science to environmental science problems [15]. Environmental science subsumes a rich tapestry of sciences, methods, and knowledge of the Earth and its climate. Environmental processes are often complex, involving non-linear interactions between biological, chemical and physical processes across a range of temporal and spatial scales [33]. The disciplines and sub-disciplines within environmental science are often fragmented and siloed, with each having their own languages, methods, conventions and approaches [16, 43, 61, 114]. There are a small number of ‘big science’ fields, that are voluminous in the quantity of data but homogeneous in instrumentation, format, content and structure and a larger number of ‘small science’ fields in which there are smaller amounts of data that exhibit greater heterogeneity and variety [17, 150]. Consequently, there is a variegated assemblage of heterogeneous structured and unstructured data—not always adequately defined or explained—at different scales and different levels of veracity, stored in field notes, local records offices, and archives as well as online repositories [33, 40, 66, 141, 144].

Given this, accessing data or outputs can be difficult to navigate, time consuming and, often ineffective [33, 90]. Even with significant

expertise, environmental data is not necessarily straightforward. Something as outwardly simple and ‘everyday’ such as soil is incredibly complex, uncertain, multi-scale and with impacts on multiple stakeholders and areas, e.g. food security, carbon capture, and protection of biodiversity [86]. Similarly, sensors—often assumed to be inherently objective—can be biased depending on their placement and other contextual factors [81, 141, 148]. Investigating global plant data, Meyer et al. [108] found that there are many uncertainties present within the data, and that any filtering done to exclude these uncertainties would drastically affect the volume of usable data. They conclude that an *opportunity to scrutinise data* is needed in order to be able to make sufficient use of it [108]. In a similar vein, Pescott et al. [122] state that there is a wide variation in the quality of environmental data, advocating for *improved meta-data that records assumptions and decisions made with the data pipeline*.

An opportunity to scrutinise data and models alongside improved supplementary information are therefore essential mechanisms to enhance scientific research, and even more so when considering the growing emphasis on multi-disciplinarity and hence the increase in non-domain experts and non-experts potentially searching for and using environmental data [44]. This is all the more important when we view data through a socio-technical lens [42]. Data is aggregated, processed and then circulated through different sites and practices [8, 58]. Data is not innately neutral, there are a multitude of judgements, biases, limitations, assumptions and uncertainties within [27, 42, 99, 134, 151]. Given these complexities, the application of data science to any area is challenging [102, 121]. However, for environmental data science in particular, these challenges are compounded [15]. With regards to environmental data science, then, trust is valuable to reduce—but not completely eliminate—complexity [37, 88, 94, 98, 111]. Particularly for stakeholders who lack familiarity or face uncertainty [87, 124, 130]. Systems that can therefore encapsulate these components, enabling collaboration, investigation, and importantly, support stakeholders to place or refuse trust appropriately are worthy of attention.

### 2.2 Collaborative Virtual Research Environments

To address the pressing environmental challenges of the planet, a holistic approach that consists of collaborative, open and reproducible scientific work is needed [15, 38, 61, 114, 153]. Transdisciplinary research, transcending any one field or approach by synthesising different types of knowledge is a promising method for addressing complex, real-world environmental problems [72, 103]. This includes not only ‘experts’ in the traditional sense, but also communities outside of academic institutions [7, 44, 89, 97]. Interdisciplinary dialogue and practice is needed, but can be problematic and a site for ‘science friction’ [9, 14, 41]. This has led to calls for the development of infrastructure to foster collaboration and bridge the gaps between scientific disciplines, policymakers, external organisations and agencies, the media, and publics [15, 38, 97, 103, 144, 152].

Virtual Labs represent a means to provide this infrastructure [25, 71, 86]. Virtual Labs are online research environments capable of providing access to datasets, analysis and visualisations for a range of stakeholders [43]. Virtual research environments are

conceptual and thus differ with each instantiation. Within environmental data science, there are realisations of the concept, such as DataLabs [71], the Biodiversity and Climate Change Virtual Laboratory [63], and the Environmental Virtual Observatory [43]. Despite the differences in how Virtual Labs may be actualised, there are some commonalities. They are cloud-based platforms enabling users to combine data and modelling software through a Web browser. Virtual Labs are differentiated from data repositories, which typically focus on meta-data, provenance, classification and standardisation [17]. These environments thus incorporate the traditional features of repositories, but also focus on tools and services, for instance conducting experiments or including elements of social networks [10, 63, 156].

Virtual research environments are typically designed with one specific community in mind, but are available to other stakeholders. By virtue of being accessible through a Web browser, Virtual Labs are technically accessible to all potential users [43]. However, open access does not equate to being usable or useful in reality. To analyse data and understand modelling functions and outputs, specific skills and knowledge may be required [60]—thus limiting the use of these environments dependent on ability. When we think about legislators or members of the public, possessing the technical skill to use new technologies or traverse a large amount of evidence should not be taken as a given [56, 140], and thus must be taken in consideration when designing virtual environments [60]. Implementations that can be taken to accommodate different stakeholders include environments within the Virtual Lab specifically tailored to skill levels or user experience [60, 71] and the inclusion of glossaries and background information [63].

There are various aims and goals for virtual environments, such as reproducibility and reusability [6, 10, 24, 30, 32, 44, 57, 100]; cross-disciplinarity and collaboration [6, 24, 44, 57, 71, 115, 156]; accessibility and understanding [7, 44]; transparency [30, 32, 44, 71, 115]; improved documentation and meta-data [158]; interactivity and usability [25, 44, 57]; interoperability and integration [6]; scalability [6, 44]; and trust [7, 32, 71, 115, 141, 156]. These are realised through various functional and socio-technical properties. For instance, flexibility enables users to define and tailor according to their specific needs. This is necessary in order to address the fact that each discipline and project is different, and has different norms and needs [25, 101, 128, 150]. Virtual Labs represent a space which allows for improved data and software integration, which can often be difficult and time-consuming on the one hand, and lead to a poor understanding and decision-making on the other [6, 64, 128, 144, 156].

However in addition to a purely technical focus, Virtual Labs are *socio-technical*, providing space for communication and collaboration across organisations and geographical locations [25, 57, 97, 118, 121, 130]. Many of these scientific communities work in isolation and so technology that can transcend the spaces between them whilst also including other stakeholders can act as a source of connection and coordination [43, 102, 121, 158]. This viewpoint enables us to think about properties from different angles. For instance, usability can be seen as a purely functional or technical property [6], or as a socio-technical property that should be defined and enacted in a community-driven way rather than as a top-down implementation [25]. Features and properties are also interlinked with others, for instance, usability can be connected to flexibility,

collaboration, and accessibility. Virtual Labs can therefore present a means to connect these standalone features.

### 2.3 The Challenge of Trust

This approach, however, is not without obstacles: “a challenge common to all emerging collaborative environments that promote open science and the rapid exchange of experimental and pre-publication data and methods is one of trust” [10, p. 609]. Taking into account multiple stakeholders and their concerns imports a consideration of a multiplicity of trusts within the design of Virtual Labs [23, 74, 115, 141, 145]. For instance, the promotion of trust is related to: the adoption and use of Virtual Labs; the research outputs included within the lab itself; and other stakeholders using the Virtual Lab.

Whilst being theorised in a myriad of different ways, one element that is common to studies of trust is an emphasis on *evidence of trustworthiness* [22, 28, 29, 84, 91, 94, 102, 104, 107, 133, 142]. Trust is often interlinked with interpretation and quality, and is commonly associated with the inclusion of provenance information as evidence of trustworthiness and to aid interpretation and hence foster trust for data and model consumers [10, 17, 32, 100, 115, 156]. Virtual Labs, as a result of their specific features and properties, allow for much needed and improved documentation, provenance and supplementary information [158]. However, given that these environments are “aimed at serving a variety of user groups that will include both method development and scientific discovery and support for decision making in environmental change, it is critical that results presented are meaningful for all stakeholders involved” [71, p. 11].

While there are contexts in which these functional and socio-technical properties have been shown to promote trust, in what follows we explore the extent to which these properties can be said to *afford trust* in Virtual Labs with treatment of multiple stakeholders. To be able to do so, we must briefly explain our particular orientation to affordances and define what we are calling ‘trust affordances’.

## 3 APPROACH

The concept of affordances is useful to guide the design of technology, capturing the relationships and interactions between users and an artefact [51, 65, 76]. Gibson [53, 54] coined the term affordances, referring to a relation (action possibility) between an environment and an organism. Gibson’s conception focused on visual, direct perception on cues for affordances, that do not change with needs and goals of an actor. This interpretation is limited when applied to technology, particularly because not all technological affordances are directly perceptible [77, 105, 146]. Consequently, affordances were adapted and adopted within HCI [51, 112, 113]. In this line of thought, affordances are processed in the brain rather than directly picked up; affordances are reliant on knowledge, culture and experience, rather than existing independently; and can be learnt whereas for Gibson they cannot be [5, 106, 116, 143, 145].

Whilst Gibson focused on utility and the possible actions afforded, the focus within HCI has been on use and usability leading to an over-emphasis on user interfaces [3, 65, 106, 116, 135, 145, 147].

When designing trustworthy systems, “the role of systems designers and researchers is thus not one of solely increasing the functionality and usability of the systems that are used to transact or communicate, but to design them in such a way that they support trustworthy action and—based on that—well-placed trust” [131, p. 383]. With this in mind, we take a view of affordances as designed elements that affect not just the human-computer dynamics (how any given user interacts with the tool) but on broader social dynamics unfolding through the technology’s mediation.

### 3.1 Trust Affordances

To develop ‘trust affordances’<sup>1</sup> we look to approaches that go beyond a limited focus on usability to those that view affordances as user- and context-specific [3, 93, 138, 146]. One means of doing this is by utilising structuration theory developed by Giddens [55]. Giddens argues that there is a duality of structure comprised of human agency and social structure. Social systems are produced and reproduced recursively between the two. Structuration theory is beneficial for investigating groups or organisations and has been applied in work contexts [145, 146] and in online communities [75]. In the same way that social structures are recursively reproduced and may be maintained or changed, every engagement with technology may see a different structure being enacted, and therefore affordances may also be reconstructed over time [120, 146]. Technology is both a product of human actions within a specific socio-cultural context and a medium for actions due to its material and social properties inscribed into them by designers [145]. Additionally, structuration theory allows us to look at technology being used in ways that the designers had not envisioned, either through appropriation or enactment [62, 120]. In this way, users do not use a technology as it was designed, and the structures within it, but may ignore, work around, or invent new properties themselves [120, 145]. This lens allows us to understand what the technology affords users and how ways of using a technology are instantiated in users’ practice within different contexts and over time [62, 116, 132, 145, 146]. It also allows us to think about how affordances can play a different role in the design of attributes of the same artefact for multiple different users [65, 74, 145].

Recent conceptualisations of affordances, such as social affordances [19, 20], collaborative affordances [5], and organisational affordances [145] are useful to us in developing ‘trust affordances’. For instance, the social and cultural context of users can directly influence how affordances are perceived [5, 20]. And how the conditions of technology, culture, power and interpretation have a combined influence on the emergence of affordances; and the emergence of a certain affordance can reinforce the conditions or, in some cases, change them [145]. Taken together, these approaches emphasise the impact of learning and change on affordances and that they may need to be discovered from practice, and that this practice is always situated in a specific and often complex socio-cultural context [21, 48, 73, 74, 145]. The relationship between the user, the context and the artefact thus affords a particular action or activity [18, 145]. Individual users can employ a specific technology for different goals and have agency in how they use it [36, 48, 135].

<sup>1</sup>We note that the phrase ‘trust affordances’ does appear within literature (cf. [32, 49, 85, 137]) but has not been formally defined in its own right.

We are therefore cognizant of the fact that however we might define trust affordances, it is dependent on them being seen as such by users.

To construct our definition we utilise the criteria for affordances as set out by Evans et al. [47]. An affordance is something that is: not solely affiliated with a platform or a feature of a platform; not an outcome; but does vary (for instance, anonymity, persistence, searchability, and visibility are affordances but privacy is not as it is an outcome) [47]. In our specific context, trust and trustworthiness is the goal or outcome, and there are affordances that reflect the means through which this can be attained.

***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.*

We proceed in the next section to explore whether and how/not commonly presumed trust-supporting properties of Virtual Labs meet this criteria for a trust affordance.

## 4 APPLICATION OF TRUST AFFORDANCES

Virtual Labs enable certain trust affordances that can shape but not dictate how people interact with Virtual Labs, with data and models, and with other users. Properties of Virtual Labs introduce action possibilities and lend themselves to specific affordances. For instance, the properties of usability, accessibility and flexibility draw upon ‘functional affordances’ (usefulness) and ‘cognitive affordances’ (understandability) [65]. Other properties, such as transparency and features such as provenance are often assumed to be trust-supporting properties—and are assumed to be, themselves, affordances. But they do not determine how one interacts with a system, or with others who are also part of the system ecology. In truth, they *imply* certain affordances, which as we show below, do not necessarily or automatically lend themselves to promoting trust.

Before launching into our examination of how provenance and transparency relate to trust in the context of Virtual Labs, we must briefly expand upon the point made within Sect. 2.3 that they are both delivered via the mechanism of documentation. Documentation is thought to afford trust through visibility, inspectability, and the provision of contextual information for stakeholders to use relative to their needs and goals, e.g. functional affordances [65]. Documentation is, at least in principle, beneficial for providing baseline evidence of trustworthiness. It can be used to ascertain whether a dataset is appropriate for a users’ needs. This is dependent not only on the quality of information within any form of documentation, but as we will show, on how that documentation is designed. To put it succinctly, there is a difference between information and *useful* information.

### 4.1 Transparency

Transparency is thought to be essential for reproducible science [139], to enable opportunities for scrutiny [115], and to improve access to scientific evidence [71]. But transparency comes in degrees: if conceived as a metaphorical ‘window’ onto information, the glass itself can range from invisible to semi-opaque, and the frame can

be the full width of the wall or a narrow aperture. The reason to carefully vary these parameters is, counter-intuitively, to optimise visibility of *relevant* information (cf. [45]). So-called ‘transparency’ has the capacity to contribute to opacity when relevant information is buried in an abundance of irrelevant information, or when data is provided in an undigested and unintelligible form. Looking for detailed evidence can be demanding, hard and time-consuming—particularly when this evidence is used to judge trustworthiness [118]. Too little detail does not provide sufficient evidence of trustworthiness, but neither does too much [13, 35, 123, 125, 154]. Transparency is well-placed to promote trust, it would appear, only when it lessens rather than increases the cognitive burden on the given user to interpret and understand information [67, 79, 126, 127]—specifically, when users are able to see all relevant data to them and their goals, but no more than this [82].

Given the existing literature exploring how to afford visibility through shades of transparency (i.e. translucency), we will focus our attention in what follows on lesser noted affordances that relate to realising transparency.

**4.1.1 Scrutability.** The ability to scrutinise, or the capacity to understand through examination, is often implied by transparency [95]. Within a Virtual Lab, scrutability is concerned with the amount of the system and its features that are shown to users. This requires careful forethought and a balance between complete openness on one side of the spectrum, and complete closure at the other end. Visibility and lack of privacy afforded by transparency may be unwanted affordances for some stakeholders [46]. Science, by its very nature uncertain and experimental, necessitates that certain elements of research are conducted behind closed doors; as in real life, so too is it the case online. Credential-based access [46, 71] is one means of providing security and anonymity for on-going scientific work and social connections, thus not being completely transparent. It also allows scientific researchers and others working on multi-disciplinary projects to have control over what is published within the Virtual Lab.

On the other hand, designers of these systems might seek to provide a certain amount/kind of evidence to promote trust in both the scientific research at hand, in other Virtual Lab users, and in the Virtual Lab itself. This being the case, one would need to consider how specific elements are published and made public, and further, how users can search for or request additional information should they require. Rather than aiming to be transparent—which is not always suitable or ideal—a more appropriate ambition might be to foster *a spirit of openness*. Openness connotes an ability to engage skeptically with a system, obtain additional information, and to scrutinise various elements in determining trustworthiness. The provision of documentation may already satisfy users, but when it does not, other scrutability-supporting features are needed. The ability to ask questions and receive reliable answers can enable stakeholders to withhold or place trust [23]. This raises certain questions: what questions do people need answers to and what evidence (or components) would serve to answer these questions? More specifically, how can we facilitate the actions of probing and conversing? And fundamentally, how can designers of these systems uncover what users’ needs are, given that they may not be able to articulate them in technical terms?

The ability to scrutinise is necessary for Virtual Lab users, such as members of the public and other groups of external stakeholders—but only up to a certain point. This affordance requires practical constraints to issues raised in our discussion of translucency, namely the tendency to overwhelm users and undermine trust in the process. It is easy to overestimate the amount the amount of times users are likely to spend scrutinising, which makes it all the more important that they can find answers to their specific questions easily. This is not easily resolved given that users inevitably seek out different confirmatory evidence of trustworthiness.

**4.1.2 Abstractability.** Virtual research environments have the ability to “provide very different information types from a single data set, and to present them in different ways that are accessible and convenient for the intended user” [44, p. 3]. It is relatively simple to place all information pertaining to a particular dataset or model online, the challenge rests in ensuring that it is accessible, intelligible, and assessable [118]. In the co-creation of an environmental virtual observatory for flooding, Elkhatib et al. [44] conducted workshops between various types of stakeholders (environmental scientists; policy makers; local communities directly involved such as farmers; and the general public) to understand what each different group required. Within their work, they opted to maintain a user experience that focused on the environmental questions asked, rather than the technical questions about how data was collected or how a model was calibrated, noting that access to these areas is available if the user wants to, but this is not immediately presented [44]. DataLabs [71] also enables a range of abstractions, allowing users access to the same project but dependent on their ability, e.g. users can work with raw code at one end of the spectrum through to graphical user interfaces on the other. In addition to this, DataLabs includes the use of notebooks, which can take the form of tutorials or contextual narratives that afford accessibility and understanding [71]. This flexibility and multimodality of information affords users the same access to the same data and models, and does not limit or exclude those who may lack programming skills, for instance.

Key to the utility of abstractability, it would seem, is first determining how this range of information is perceived by those it is aimed at. This approach aims to codify specific types of knowledge that are gained through experience, however this is difficult to transfer to others [92]. Who creates these abstractions? Do they convey what they are meant to convey? Do they enable understanding? Which formats of abstraction are suited to whom? How and where is this abstracted information used and how can it be restored if and when it becomes relevant later on? In short, what should be abstracted out, for whom, and how do we know it has led people to the information they need to form well-placed trust?

**4.1.3 Reciprocity.** In naming this affordance we harken to the literature on reciprocal design—here, specifically, highlighting the need for documentation to be tailored to the individual needs of recipients. Annotation, enabling stakeholders to add to a document and hence modify it, is one way of achieving this [65]. Annotation can reduce the epistemic load placed onto creators and reduce the expertise gap between creators and stakeholders [34, 128]. Annotation further affords trust through collaboration, e.g. a shared sense of ownership, a sense of adding value or contributing, and contributes to the holistic research vision for environmental data

science. Annotation by other users, such as those who have used the data and left comments on it, may afford trust in determining appropriateness, completeness and accuracy. Annotation is also a way to pose questions or solve uncertainties. Documentation is predominantly created by those who collected or aggregated the data or who ran and tuned a model. We should not assume that those same people can inherently understand all questions held by others.

Conversation enables the ability to scope, rephrase or ask additional questions, affording mutuality and responsiveness [59, 67, 110]. Being able to ask questions and receive reliable answers are well-placed in promoting trust, aligning trustors and trustees expectations [23, 115]. For instance, a particular user of a dataset can determine the reasoning of the data producer and how this relates to the appropriateness of said dataset to their own specific needs. When the intelligibility of documentation cannot be guaranteed, features to support additional dialogue are worthy of investment to promote trust.

How might these features afford trust and for whom? Reciprocity is afforded through communication, the sharing of dialogue, and the ability to build relationships. Trust is individual- and context-specific [102]. Therefore we cannot pre-meditatively completely compile all evidence of trustworthiness for any given stakeholder. We can generalise, and provide the information we think might be meaningful, but without a mechanism to address any questions, the promotion of trust is hindered. Annotation goes some way in addressing this, with comments left and attached to documents. But annotation and tagging can be problematic in its own right (cf. [75]). In addition, then, channels of communication, are required. Enabling various stakeholders to gain access to the information that they need in order to be able to place or refuse trust appropriately. But this too can be fraught with design decisions to be made and affordances produced. Person-to-person private messaging enables direct questions to be asked and answered, but are ephemeral and contained within the specific interaction [45]. On the other hand, discussion boards enable conversation to be broadcast to others. This prevents the same question being asked numerous times, but also introduces the problems of not knowing the audience and navigating messy threads [18, 45]. Within our reflections, it seems evident that some form of communication is required, but the specific instantiation of this is to be determined. We are aware that channels of communication are already pre-existent for scientists and academic researchers, e.g. through personal acquaintances or by virtue of having an institutional affiliation. We believe that these affordances should be expanded to a wider demographic of Virtual Lab users to ask, answer, and discuss.

## 4.2 Provenance

Provenance is slated as critical for assessing authenticity [139], maintaining integrity in scientific results [100], and (linking back to transparency) enabling the transparency and reproducibility of scientific results [30]. But as we saw with transparency, ‘provenance’ varies: basic forms describe the source and lineage of data; complex forms of provenance describe production processes and influences such as assumptions or bias [26, 31, 139]. Complex forms

may also be referred to as supplementary information. Spiekermann et al. [139] delineate between coarse-grained provenance, which captures relationships but not complete insight, and fine-grained provenance, which provides a comprehensive account of all information pertaining to a dataset or model. Therefore what is stated as ‘provenance’ is a specific form of provenance in every case. Cheney et al. [26] also characterise different types of provenance within a database: ‘where’ (e.g. where data has come from?), ‘how’ (e.g. how was an output produced?), and ‘why’ (e.g. why was an output produced?). Documents can have different values—for instance, evidential value relates to structure, procedures, and proving ownership; and informational value refers to the contents of the document for reference, contemplation and research [67]. Thus when provenance is hailed as a critical foundation crucial to scientists, which provenance are we talking about? Coarse-grained, ‘where’ evidential provenance, or fine-grained, ‘why’ informational provenance? The varying granularity of provenance connotes different affordances. The former type of provenance affords insight into the origins of a particular dataset, model, or outcome. It is not transparent, capturing the relationships between items but excluding the complete derivation [139]. In our context, this type of provenance affords trust when there are issues of legality or ownership, as it can be used to ascertain basic facts about a sensor and its recordings, for example [67, 141]. The latter provenance, which is fine-grained and informational provides much deeper insight into these relationships. For instance within scientific research, the onus is on the contents of any documentation or provenance information, including any supplementary details which provide value [67, 139]. The granularity of provenance is not dichotomous or mutually exclusive, but enables us to consider the affordances between the types. Coarse-grained may be sufficient for a stakeholder who wants to see where something has come from, or been used in but for whom evidence of trustworthiness is not reliant upon seeing all details. Whereas fine-grained may be useful to a modeller, who needs a comprehensive account of what has been done, how, and why in order to have trust in the source and appropriately conduct their research.

Trust is a judgement based upon available evidence, but may also rely on many other elements in addition [102]. Approaches that focus on ‘where’ questions and evidential value only go some way in affording trust. Evidence of trustworthiness is not prescriptive, and scientists may rely on approximate evidence and heuristics in addition to provenance information, e.g. reputation; comparison with standards; or appropriateness for their needs [50, 80, 118, 149]. The validity of data is related not only to lineage, but also supplementary information, such as context and instrumentation [102, 141]. To assess and use data sufficiently, data consumers may need to look beneath mere provenance records to look at the assumptions or to answer ‘how’ and ‘why’ questions [8, 26].

**4.2.1 Verifiability.** Trustworthy data is data that is verifiably accurate, that has evidence of integrity, authenticity and reliability [81, 100, 136]. It is not purely about the data journey, but the accuracy and completeness of the source and the provenance information itself [56, 100]. Verifiability can be achieved through identity, or can be achieved through things such as distributed ledger technology.

Knowing in whom you are placing your trust enables an individual to make interpersonally rooted trust assumptions; and conversely when identity is unknown, assumptions of trustworthiness are based in more impersonal dimensions [28]. Hoffman et al. [69] propose the use of blockchain for academic published papers which will enable verification of collaborators' identities. Use of distributed ledger technology would also enable data consumers to know that the record is authentic in what it purports to be and is free from corruption or tampering [136]. However this has the potential to be troublesome for sensitive or proprietary materials [30, 155]. Thornton et al. [141] note that the use of blockchain within academic research could lead to a paradox where evidence of trustworthiness is provided for some consumers, such as those who seek quality assurance, but simultaneously lead to disengagement for others who prioritise the protection of business interests and intellectual property.

Verifiability is therefore a two-sided affordance, arising from a tension between privacy and visibility [46]. This tension is exacerbated if we utilise technology, such as blockchain, to retain persistence and immutability. But then in some cases retaining information about one's identity within a system can be beneficial in supporting trustworthy behaviours (e.g. providing awareness of and accountability for one's actions [45], retaining a memory of past behaviours in the form of a reputation [28]). Potentially problematically, in a system with multiple stakeholders, it may be that for some users preserving identity is key to being able to verify and trust, and for others it inhibits collaboration and reduces trust. Balancing these issues, then, and deciding on what level of identity to present in order to promote well-placed trust is highly context-specific and a system may need to be flexible to accommodate individual users' preferences.

**4.2.2 Persistence.** Within our context, persistence relates to trust that the data or models will remain in the same content and location, unless a new version exists (cf. [29]). Effort is required to develop a persistent infrastructure for artefacts ranging from datasets to models and supplementary information about them [9]. On the future of data and its quality, Lord et al. [96] ask whether all data should be kept and what issues may arise from decisions undertaken. The velocity of data is increasing exponentially, and the veracity of data (in particular environmental science) is often questionable. Given this, decisions need to be made to increase computing capacity to store *all* data regardless of its relevancy for future research, e.g. potential inaccuracies or incompleteness—or do we selectively filter data before we commit it to a repository? Both approaches may be unsatisfactory. There are only so many bytes of data, model runs, and outputs we can store. But on the other hand, we cannot necessarily decide what may be useful to someone in the future, and may inadvertently privilege specific types of information over others in doing so.

If persistence is deemed important to trust within the context of a Virtual Lab, there are design challenges related to curation, archival, and preservation [96]. Not only does the data itself need preserving in many cases, but so too does associated meta-data, so as to facilitate *trustworthy* reuse of the data (cf. [10, 156]). But what meta-data is useful to a given stakeholder, and what meta-data a given stakeholder is capable of collecting, may differ greatly; in

which case, how are these trade-offs and potential inconsistencies resolved?

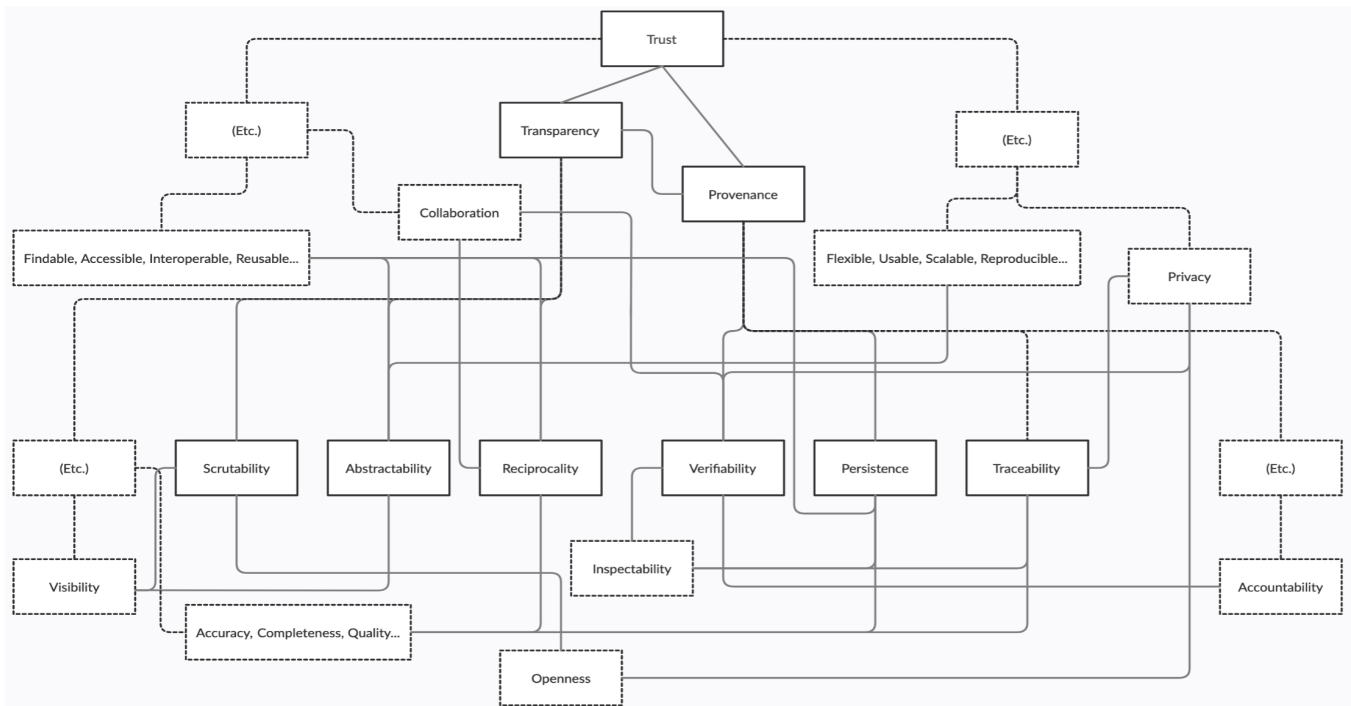
**4.2.3 Traceability.** Environmental data may have tangled heredity and ownership, which may entail more sources than a user is able to assess at one time [81, 102]. It can be challenging for any user, regardless of expertise, to assess the quality of aggregated environmental data [4, 15, 50]. Barclay [4] suggests that a gateway (in our context, a Virtual Lab) can dually: provide traceability, e.g. origins and reuse of data; and importantly, assist with navigation through this complexity. For instance, myExperiment [10, 32] facilitate an infrastructure to 'bundle' essential information pertaining to scientific research into workflows. However, workflows can break [157], and from our own experience workflows can be limiting: on the inputs and outputs, data accessibility, data sharing, and on the stakeholders that can use workflows, e.g. external project partners with limited technical ability.

Traceability, whilst often thought of in terms of looking back, can also include forward tracing [81]. Scientists can often be reticent in sharing data because of a fear that their data will be taken and used inappropriately or for erroneous science [141]. The work on DataLabs advocates for strong supporting mechanisms for trust when sharing data [71]. Once more, however, this approach is two-sided. Trust may be afforded for scientists with evidence of the propagation of their data, but this may be an unwanted affordance for commercial users of Virtual Labs, who do not want to share or promote which data they use within their analysis and want to protect intellectual property [141]. Any solution to this would require some type of constraint, which Clark [28] argues, nonetheless cannot *force* people to behave in a trustworthy way. In the face of increasing calls for open data, traceability may be suited to some forms of data or research, e.g. high profile or contentious issues, but not suited to the majority of research, for which backwards and forwards traceability would be problematic.

### 4.3 Making Connections

Following the criteria set out by Evans et al. [47], neither transparency nor provenance are in themselves affordances, but relate to a number of affordances which can—for some stakeholders in some contexts—support trust. We have knowingly made a somewhat unnatural separation between transparency and provenance, as they are intimately connected through these various other affordances. We have tried to demonstrate that affordances may be two- or multi-sided [74]: the affordances associated with transparency and/or provenance can be positive for some users but may have negative implications for others. Furthermore, whilst we have expanded upon six trust affordances, which can all contribute to trust, we should not assume that they afford trust in *all* instances or *individually*. For instance, reciprocity and scrutability combined may afford trust, where neither does in and of itself in a particular scenario.

A mind map (Figure 1) captures part of the complexity in designing for trust. As messy a picture as this already is, we note that it is incomplete. We have not fully unpacked all of the affordances related to transparency and provenance (if there even is a bottom to this well), and these are hardly the only considerations in designing for trust. At a minimum, the map shows that simple recipes



**Figure 1: Mind map, illustrative of interconnections and tensions within and between affordances relating to transparency and provenance. Dotted lines indicate concepts we have made connections to without fully unpacking.**

(e.g. FAIR represents data management that is Findable, Accessible, Interoperable, and Reusable [71, 78]) inevitably elide myriad important design considerations which impact on the resultant trust.

We are not arguing that transparency and provenance are antagonistic to trust—clearly they can be appropriate high-level design goals. But promoting trust is clearly not as simple as adding a bit of transparency and a bit of provenance. The affordances that feed into both are not clear-cut themselves and evidently require careful consideration of potentially conflicting multi-stakeholder requirements. It is important to clearly define sub-goals that not only support transparency and provenance *but also support trust* (remembering that one does not guarantee the other), drawing on various affordances with appreciation of their interdependencies and tensions. Most importantly, we are by no means suggesting designing for trust is futile; rather, we are trying to foster a greater sensibility regarding the delicacy of designing for trust, and the importance of really understanding the requirements of the different users of the system.

## 5 DISCUSSION

The design process is fraught with decisions to be made and competing interests to be balanced, which as we have shown, are not necessarily as simple as they initially appear. Trust is recognised as being key to the success of socio-technical systems; and because trust is notoriously difficult to retrofit, how a system will foster trust needs to be considered early and throughout the design process [82]. By exploring how properties designed into Virtual Labs either

do or do not afford trust, we have illustrated the need for a nuanced (rather than prescriptive) ‘trustworthy by design’ approach that accounts for how system affordances relate to the specific trust needs of individual stakeholders. As we have shown, certain features of Virtual Labs both afford trust and *do not* afford trust, depending on the context. The broader lesson in this is to be careful in proclaiming a *natural* relationship between any particular system property and trust.

### 5.1 Trust Affordances in AI Literature

Trust has garnered significant attention within the computing community, particularly in the Artificial Intelligence (AI) field. Prevalent conceptions of trusted AI invoke notions of fairness, accountability, transparency and provenance. As we have seen above, implementing technology and/or technological artefacts in a way that affords trust to multiple stakeholders requires careful consideration and is not an easy task to undertake.

Transparency in particular is leaned on heavily in the trusted AI community, increasingly instantiated through forms of documentation [1, 11, 52, 70, 109]. As such, there is much from our discussion of Virtual Labs that can be applied to trusted AI more generally. AI documentation across the board is designed (implicitly) for scrutability, enabling stakeholders to examine relevant facts of the AI’s development. There has been much interest in seeking to *enable more dialogic scrutability*, for example providing explanations (whether in documentation or via other means) in the form of counterfactuals [110], which raises similar issues to the ones we



have here about the challenges in predicting the questions individuals might have. Reciprocity is possibly less universal though clearly features in IBM FactSheets, which are tailored to specific documentation consumers [129].

It seems inevitable that some of the subtleties of the various affordances subsumed within transparency arose in designing these forms of documentation and influenced the particular form this transparency took. Our point is that it is important not to lose that subtlety when justifying documentation in terms of how transparency contributes to trust in AI. If and when documentation succeeds in doing so, it is because its designers have attended to balancing myriad interconnecting considerations that impact on trust. Where they fail to promote trust, it is likely due to insufficient attention to specific multi-stakeholder trust needs. Transparency should, therefore, always be approached within AI as a ‘problematic ideal’ [121] that requires great specificity regarding the trust needs of relevant users and reflection on how designed aspects of systems might contribute to (or undermine) the fulfilment of these needs.

## 5.2 Scaling Trust

Many if not all of the affordances described in this paper speak to *interpersonal* trust, which requires careful nurturing in ways that do not necessarily scale as systems become more complex. The Virtual Labs we have described are relatively small communities: the infrastructure provided can enable the nurturing of trust at a project- or team-level. But can these environments scale any further? At some point, it becomes impractical to individually assess trustworthiness and one instead must place their trust in the system itself as having inbuilt mechanisms for assuring trustworthiness (cf. [83]). Interestingly, this suggests that Virtual Labs, if scaled up, might entail a rather different set of affordances related to *impersonal* (or *system*) trust.

Our exploration of transparency and provenance affordances—our overwhelmingly complex mind map (Figure 1) in particular—also raise more general questions about how much effort is reasonable to expend in support of trust given how much is involved in developing systems that enable trust to thrive. Trust really doesn’t scale well. For example, documentation is costly to create and maintain, requiring a high level of manual input [1, 68, 70], as are features that enable provenance of scientific data [24, 141]. On balance, then, how important is trust? How do system developers make decisions about how much trust is enough trust? And how do we approach the design of a system given these considerations? The trouble is, problems stemming from a lack of trust are not easily fixed once they appear—hence the emphasis in the literature on a trustworthy by design approach [82]. This is not something we can answer here, but pose it as a question for ongoing deliberation by the research community.

## 5.3 Further Trust Affordances

Whilst not the focus of our paper, we note that given the inherent collaborative nature of environmental data science and the potential collaborative affordances offered by Virtual Labs, we would be remiss to exclude a consideration of trust affordances pertaining to collaboration. These dynamics are complex, but as we have alluded to in Figure 1, are present within design considerations of

virtual research environments, e.g. collaboration is linked to both reciprocity and verifiability. Whilst we have not unpacked further additional trust affordances of collaborative research environments within our research, this area is one amongst others that we hope to build upon in future work.

## 6 CONCLUSION

In this paper we have introduced and defined trust affordances, characteristics of a technology to promote trust through the provision of evidence of (dis)trustworthiness. Our effort was driven by our own experiences in designing features within a virtual research lab for environmental data science—specifically, the realisation that when trust is discussed, there are certain (often implicit) features which are assumed to unwaveringly support trust... and which end up being far more complicated in their implementation. In particular, provenance and transparency are two such features often presumed to provide authentic, informative evidence of trustworthiness. However when we uncover the complexities of designing these properties into systems, we find that there are no blueprints or maps to guide us in enabling said design, and moreover, there is no clarity on which features actually equate to trust. Given this, we argue that we should be analytic within the design process, particularly of the prevailing assumption that designing a system consists of a sprinkle of one element here, and a dash of another element there. In reality, a critical understanding of the social realities in which any system is situated is crucial. In the design of Virtual Labs, we have learnt that provenance and transparency are useful, with the caveat that these features need to be articulated in multiple ways in order to contribute to trust and what’s more we must be cognizant of the fact that when a system is designed in such a way as to promote trust for one stakeholder, the same practice may not be promoting trust for another. With regards to the wider community utilising documentation for the promotion of trust, in particular data science and AI, we believe that these reflections can be beneficial to guide the design of systems and features within these systems. There are different forms of provenance and transparency that can be taken, and not all roads lead to trust. The use of trust affordances has facilitated the consideration of features and systems as-a-whole from multiple perspectives and contexts, ones which can easily be glossed over. We hope that our work can enable greater emphasis on the differing stakeholders within each context and inclusion of additional mechanisms to promote trust. Ultimately, we want to show that despite the numerous shades of grey within trust, it is only when we embrace this that we can design in a trustworthy manner.

## ACKNOWLEDGMENTS

We would like to thank the reviewers for their feedback. This research was supported in part by EPSRC (grant number: EP/R512564/1).

## REFERENCES

- [1] Matthew Arnold, Rachel K. E. Bellamy, Michael Hind, Stephanie Houde, Sameep Mehta, Aleksandra Mojsilovic, Ravi Nair, Karthikeyan Natesan Ramamurthy, Alexandra Olteanu, David Piorkowski, Darrell Reimer, John T. Richards, Jason Tsay, and Kush R. Varshney. 2019. FactSheets: Increasing trust in AI services through supplier’s declarations of conformity. *IBM J. Res. Dev.* 63, 4/5 (2019), 6:1–6:13. <https://doi.org/10.1147/JRD.2019.2942288>

- [2] Maryam Ashoori and Justin D. Weisz. 2019. In AI We Trust? Factors That Influence Trustworthiness of AI-infused Decision-Making Processes. (2019). arXiv:1912.02675
- [3] Klaus B. Børntsen and Johan Trettkvik. 2002. An Activity Theory Approach to Affordance. In *Proceedings of the Second Nordic Conference on Human-Computer Interaction* (Aarhus, Denmark) (*NordiCHI '02*). Association for Computing Machinery, New York, NY, USA, 51–60. <https://doi.org/10.1145/572020.572028>
- [4] Iain Barclay, Alun D. Preece, Ian J. Taylor, and Dinesh C. Verma. 2019. Towards Traceability in Data Ecosystems using a Bill of Materials Model. (2019). arXiv:1904.04253
- [5] Jakob E. Bardram and Steven Houben. 2018. Collaborative Affordances of Medical Records. *Comput. Supported Coop. Work* 27, 1 (2018), 1–36. <https://doi.org/10.1007/s10606-017-9298-5>
- [6] Michelle Barker, Silvia Delgado Olabarriaga, Nancy Wilkins-Diehr, Sandra Gesing, Daniel S. Katz, Shayan Shahand, Scott Henwood, Tristan Glatard, Keith Jeffery, Brian Corrie, Andrew Treloar, Helen Graves, Lesley Wyborn, Neil P. Chue Hong, and Alessandro Costa. 2019. The Global Impact of Science Gateways, Virtual Research Environments and Virtual Laboratories. *Future Generation Computer Systems* 95 (2019), 240–248. <https://doi.org/10.1016/j.future.2018.12.026>
- [7] Michele L. Barnes, Emmanuel Mbaru, and Nyawira Muthiga. 2019. Information access and knowledge exchange in co-managed coral reef fisheries. *Biological Conservation* 238 (2019). <https://doi.org/10.1016/j.biocon.2019.108198>
- [8] Jo Bates, Yu-Wei Lin, and Paula Goodale. 2016. Data journeys: Capturing the socio-material constitution of data objects and flows. *Big Data & Society* 3, 2 (2016). <https://doi.org/10.1177/2053951716654502>
- [9] Neil Beagrie. 2006. Digital Curation for Science, Digital Libraries, and Individuals. *International Journal of Digital Curation* 1 (2006), 3–16. <https://doi.org/10.2218/ijdc.v1i1.2>
- [10] Sean Bechhofer, Iain Buchan, David De Roure, Paolo Missier, John Ainsworth, Jiten Bhagat, Philip Couch, Don Cruickshank, Mark Delderfield, Ian Dunlop, Matthew Gamble, Danius Michaelides, Stuart Owen, David Newman, Shoaib Sufi, and Carole Goble. 2013. Why linked data is not enough for scientists. *Future Generation Computer Systems* 29, 2 (2013), 599–611. <https://doi.org/10.1016/j.future.2011.08.004>
- [11] Emily M. Bender and Batya Friedman. 2018. Data Statements for Natural Language Processing: Toward Mitigating System Bias and Enabling Better Science. *Transactions of the Association for Computational Linguistics* 6 (2018), 587–604. [https://doi.org/10.1162/tacl\\_a\\_00041](https://doi.org/10.1162/tacl_a_00041)
- [12] Mike Berners-Lee. 2019. *There Is No Planet B: A Handbook for the Make or Break Years*. Cambridge University Press, New York.
- [13] Umang Bhatt, Alice Xiang, Shubham Sharma, Adrian Weller, Ankur Taly, Yunhan Jia, Joydeep Ghosh, Ruchir Puri, José M. F. Moura, and Peter Eckersley. 2020. Explainable Machine Learning in Deployment. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) (*FAT\* '20*). Association for Computing Machinery, New York, NY, USA, 648–657. <https://doi.org/10.1145/3351095.3375624>
- [14] Gordon S. Blair, Keith Beven, Rob Lamb, Richard Bassett, Kris Cauwenberghs, Barry Hankin, Graham Dean, Neil Hunter, Liz Edwards, Vatsala Nundloll, Faiza Samreen, Will Simm, and Ross Towe. 2019. Models of Everywhere Revisited: A Technological Perspective. *Environmental Modelling & Software* 122 (2019), 104521. <https://doi.org/10.1016/j.envsoft.2019.104521>
- [15] Gordon S. Blair, Amber Leeson, John Watkins, Emma Eastoe, Susan Jarvis, and Paul J. Young. 2019. Data Science of the Natural Environment: A Research Roadmap. *Front. Environ. Sci.* 7, 121 (2019), 1–14. <https://doi.org/10.3389/fenvs.2019.00121>
- [16] Jan J. Boersema. 2009. Environmental Sciences, Sustainability, and Quality. In *Principles of Environmental Sciences*, Jan J. Boersema and Lucas Reijnders (Eds.). Springer Netherlands, Dordrecht, 3–14. [https://doi.org/10.1007/978-1-4020-9158-2\\_1](https://doi.org/10.1007/978-1-4020-9158-2_1)
- [17] Christine L. Borgman. 2015. *Big Data, Little Data, No Data: Scholarship in the Networked World*. The MIT Press, Cambridge, MA.
- [18] danah boyd. 2010. Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In *Networked Self: Identity, Community, and Culture on Social Network Sites*, Zizi Papacharissi (Ed.). Routledge, New York, Chapter 2, 39–58.
- [19] Erin Bradner. 2001. Social Affordances of Computer-Mediated Communication Technology: Understanding Adoption. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems* (Seattle, Washington) (*CHI EA '01*). Association for Computing Machinery, New York, NY, USA, 67–68. <https://doi.org/10.1145/634067.634111>
- [20] Erin Bradner, 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* (Copenhagen, Denmark). Kluwer Academic Publishers, USA, 139–158. [https://doi.org/10.1007/978-94-011-4441-4\\_8](https://doi.org/10.1007/978-94-011-4441-4_8)
- [21] Monika Bischer, Preben Mogensen, and Dan Shapiro. 2001. Spaces of practice. In *Proceedings of the Seventh European Conference on Computer Supported Cooperative Work, 16-20 September 2001, Bonn, Germany*, Wolfgang Prinz, Matthias Jarke, Yvonne Rogers, Kjeld Schmidt, and Volker Wulf (Eds.). Kluwer, 139–158. [https://doi.org/10.1007/0-306-48019-0\\_8](https://doi.org/10.1007/0-306-48019-0_8)
- [22] V. Cahill, E. Gray, J.M. Seigneur, C. Jensen, Y. Chen, B. Shand, N. Dimmock, A. Twigg, J. Bacon, C. English, W. Wagealla, S. Terzis, P. Nixon, G. Serugendo, C. Bryce, M. Carbone, K. Krukow, and M. Nielsen. 2003. Using trust for secure collaboration in uncertain environments. *IEEE Pervasive Computing* 2, 3 (2003), 52–61. <https://doi.org/10.1109/MPRV.2003.1228527>
- [23] Muffy Calder, Simon Dobson, Michael Fisher, and Julie McCann. 2018. Making Sense of the World: Framing Models for Trustworthy Sensor-Driven Systems. *Computers* 7, 4 (Nov 2018), 62. <https://doi.org/10.3390/computers7040062>
- [24] Sarah Callaghan, Steve Donegan, Sam Pepler, Mark Thorley, Nathan Cunningham, Peter Kirsch, Linda Ault, Patrick Bell, Rod Bowie, Adam Leadbetter, Roy Lowry, Gwen Moncoiffe, Kate Harrison, Ben Smith-Haddon, Anita Weatherby, and Dan Wright. 2012. Making Data a First Class Scientific Output: Data Citation and Publication by NERC’s Environmental Data Centres. *International Journal of Digital Curation* 7, 1 (2012), 107–113. <https://doi.org/10.2218/ijdc.v7i1.218>
- [25] Leonardo Candela, Donatella Castelli, and Pasquale Pagano. 2013. Virtual Research Environments: An Overview and a Research Agenda. *Data Science Journal* 12 (2013), 75–81. <https://doi.org/10.2481/dsj.GRDI-013>
- [26] James Cheney, Laura Chiticariu, and Wang-Chiew Tan. 2009. Provenance in Databases: Why, How, and Where. *Found. Trends Databases* 1, 4 (April 2009), 379–474. <https://doi.org/10.1561/19000000006>
- [27] Elizabeth F. Churchill. 2012. From Data Divination to Data-Aware Design. *Interactions* 19, 5 (2012), 10–13. <https://doi.org/10.1145/2334184.2334188>
- [28] David Clark. 2014. The Role of Trust in Cyberspace. In *Trust, Computing, and Society*, Richard H. Harper (Ed.). Cambridge University Press, Cambridge, Chapter 2, 17–37. <https://doi.org/10.1017/CBO9781139828567>
- [29] Piotr Cofta. 2007. Confidence, Trust and Identity. *BT Technology Journal* 25, 2 (2007), 173–178. <https://doi.org/10.1007/s10550-007-0042-4>
- [30] Susan B Davidson, Sanjeev Khanna, Sudeepa Roy, Julia Stoyanovich, Val Tannen, and Yi Chen. 2011. On Provenance and Privacy. In *Proceedings of the 14th International Conference on Database Theory* (Uppsala, Sweden) (*ICDT '11*). Association for Computing Machinery, New York, NY, USA, 3–10. <https://doi.org/10.1145/1938551.1938554>
- [31] Simon de Lusignan, Siaw-Teng Liaw, Paul Krause, Vasa Curcin, M Vicente, Georgios Michalakidis, Lars Agreus, Peter Leyesen, Nicola Shaw, and K Mendis. 2011. Key Concepts to Assess the Readiness of Data for International Research: Data Quality, Lineage and Provenance, Extraction and Processing Errors, Traceability, and Curation. Contribution of the IMA Primary Health Care Informatics Working Group. *Yearbook of medical informatics* 6 (2011), 112–120. <https://doi.org/10.1055/s-0038-1638748>
- [32] David De Roure, Carole Goble, Sergejs Aleksejevs, Sean Bechhofer, Jiten Bhagat, Don Cruickshank, Paul Fisher, Duncan Hull, Danius Michaelides, David Newman, Rob Procter, Yuwei Lin, and Meik Poschen. 2010. Towards Open Science: The MyExperiment Approach. *Concurr. Comput.: Pract. Exper.* 22, 17 (2010), 2335–2353. <https://doi.org/10.1002/cpe.1601>
- [33] Graham Dean, Victoria Janes Bassett, Ross Paul Towe, Vatsala Nundloll, Jessica Davies, and Gordon Shaw Blair. 2017. An Interdisciplinary approach to Finding and Using Data for Complex Environmental Modelling Problems: A Soil System Example. *EnvironInfo2017* (2017). [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](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)
- [34] Alan Dix. 2007. Designing for Appropriation. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but Not as We Know It - Volume 2* (BCS-HCI '07). BCS Learning Development Ltd., Swindon, GBR, 27–30. <https://doi.org/doi/10.5555/1531407.1531415>
- [35] Finale Doshi-Velez and Been Kim. 2017. Towards A Rigorous Science of Interpretable Machine Learning. arXiv:1702.08608
- [36] Paul Dourish, W. Keith Edwards, Anthony LaMarca, and Michael Salisbury. 1999. Presto: An Experimental Architecture for Fluid Interactive Document Spaces. *ACM Trans. Comput.-Hum. Interact.* 6, 2 (June 1999), 133–161. <https://doi.org/10.1145/319091.319099>
- [37] Natasha Dwyer, Anirban Basu, and Stephen Marsh. 2013. Reflections on Measuring the Trust Empowerment Potential of a Digital Environment. In *Trust Management VII (IFIP Advances in Information and Communication Technology, Vol. 401)*, M. Carmen Fernández Gago, Fabio Martinelli, Siani Pearson, and Isaac Agudo (Eds.). Springer, 127–135. [https://doi.org/10.1007/978-3-642-38323-6\\_9](https://doi.org/10.1007/978-3-642-38323-6_9)
- [38] Steve M. Easterbrook. 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*, Gruia-Catalin Roman and Kevin J. Sullivan (Eds.). ACM, 99–104. <https://doi.org/10.1145/1882362.1882383>
- [39] Steve M. Easterbrook. 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. <https://doi.org/10.2991/ict4s-14.2014.28>

- [40] Paul N. Edwards. 2010. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. The MIT Press, Cambridge, MA.
- [41] Paul N. Edwards, Matthew S. Mayernik, Archer L. Batcheller, Geoffrey C. Bowker, and Christine L. Borgman. 2011. Science friction: Data, metadata, and collaboration. *Social Studies of Science* 41, 5 (2011), 667–690. <https://doi.org/10.1177/0306312711413314>
- [42] Madeleine Clare Elish and danah boyd. 2018. Situating methods in the magic of Big Data and AI. *Communication Monographs* 85, 1 (2018), 57–80. <https://doi.org/10.1080/03637751.2017.1375130>
- [43] Yehia Elkhatib, 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*, Paulo Ferreira, Zbigniew Jerzak, and Peter R. Pietzuch (Eds.). ACM, 13–18. <https://doi.org/10.1145/2460756.2460759>
- [44] Yehia Elkhatib, Alastair L. Gemmill, Claudia Vitolo, Mark E. Wilkinson, Eleanor B. Mackay, Barbara J. Percy, Gordon S. Blair, and Robert J. Gurney. 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, 1404–1415. <https://doi.org/10.1109/ICDCS.2019.00140>
- [45] Thomas Erickson and Wendy A. Kellogg. 2000. Social Translucence: An Approach to Designing Systems That Support Social Processes. *ACM Trans. Comput.-Hum. Interact.* 7, 1 (March 2000), 59–83. <https://doi.org/10.1145/344949.345004>
- [46] Thomas Erickson, David N. Smith, Wendy A. Kellogg, Mark Laff, John T. Richards, and Erin Bradner. 1999. Socially Translucent Systems: Social Proxies, Persistent Conversation, and the Design of “Babble”. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Pittsburgh, Pennsylvania, USA) (CHI '99). Association for Computing Machinery, New York, NY, USA, 72–79. <https://doi.org/10.1145/302979.302997>
- [47] Sandra K. Evans, Katy E. Pearce, Jessica Vitak, and Jeffrey W. Treem. 2016. Explicating Affordances: a Conceptual Framework for Understanding Affordances in Communication Research. *Journal of Computer-Mediated Communication* 22, 1 (2016), 35–52. <https://doi.org/10.1111/jcc4.12180>
- [48] Anne-Laure Fayard and John Weeks. 2014. Affordances for Practice. *Inf. Organ.* 24, 4 (Oct. 2014), 236–249. <https://doi.org/10.1016/j.infoandorg.2014.10.001>
- [49] Ashley Fulmer and Kurt Dirks. 2018. Multilevel trust: A theoretical and practical imperative. *Journal of Trust Research* 8, 2 (2018), 137–141. <https://doi.org/10.1080/21515581.2018.1531657>
- [50] Matthew Gamble 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, New York, NY, USA. <https://doi.org/10.1145/2527031.2527048>
- [51] William W. Gaver. 1991. Technology Affordances. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New Orleans, Louisiana, USA) (CHI '91). Association for Computing Machinery, New York, NY, USA, 79–84. <https://doi.org/10.1145/108844.108856>
- [52] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Duménil, and Kate Crawford. 2018. Datasheets for Datasets. arXiv:1803.09010
- [53] James J. Gibson. 1977. The Theory of Affordances. In *Perceiving, Acting, Knowing: Toward an Ecological Psychology*, Robert Shaw and John Bransford (Eds.). Lawrence Erlbaum, Hillsdale, NJ, Chapter 3, 67–82.
- [54] James J. Gibson. 1979. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA.
- [55] Anthony Giddens. 1984. *The Constitution of Society: Outline of the Theory of Structuration*. University of California Press, Berkeley, CA.
- [56] Sarah Giest. 2017. Big Data for Policymaking: Fad or Fasttrack? *Policy Sciences* 50, 3 (2017), 367–382. <https://doi.org/10.1007/s11077-017-9293-1>
- [57] Yolanda Gil, Ewa Deelman, Mark Ellisman, Thomas Fahringer, Geoffrey Fox, Dennis Gannon, Carole Goble, Miron Livny, Luc Moreau, and Jim Myers. 2007. Examining the Challenges of Scientific Workflows. *Computer* 40, 12 (Dec. 2007), 24–32. <https://doi.org/10.1109/MC.2007.421>
- [58] Lisa Gitelman (Ed.). [n.d.]. The MIT Press, Cambridge, MA.
- [59] Alyssa Glass, Deborah L. McGuinness, and Michael Wolverton. 2008. Toward Establishing Trust in Adaptive Agents. In *Proceedings of the 13th International Conference on Intelligent User Interfaces* (Gran Canaria, Spain) (IUI '08). Association for Computing Machinery, New York, NY, USA, 227–236. <https://doi.org/10.1145/1378773.1378804>
- [60] Evgeny P. Gordov, Igor Okladnikov, Anna Ryazanova, Alexander Titov, and Yulia Gordova. 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. *IOP Conference Series: Earth and Environmental Science* 211 (dec 2018). <https://doi.org/10.1088/1755-1315/211/1/012079>
- [61] Sheila Greene, Penny J. Johnes, John P. Bloomfield, Sim M. Reaney, Russell S. Lawley, Yehia Elkhatib, Jim Freer, Nick Odoni, Christopher J. A. Macleod, and Barbara J. Percy. 2015. A geospatial framework to support integrated biogeochemical modelling in the United Kingdom. *Environmental Modelling and Software* 68 (2015), 219–232. <https://doi.org/10.1016/j.envsoft.2015.02.012>
- [62] Trisha Greenhalgh and Rob Stones. 2010. Theorising big IT programmes in healthcare: Strong structuration theory meets actor-network theory. *Social Science Medicine* 70, 9 (2010), 1285–1294. <https://doi.org/10.1016/j.socscimed.2009.12.034>
- [63] Willow Hallgren, Linda Beaumont, Andrew Bowness, Lynda Chambers, Erin Graham, Hamish Holewa, Shawn Laffan, Brendan Mackey, Henry Nix, Jeff Price, Jeremy Vanderwal, Rachel Warren, and Gerhard Weis. 2016. The Biodiversity and Climate Change Virtual Laboratory: Where ecology meets big data. *Environmental Modelling Software* 76 (2016), 182–186. <https://doi.org/10.1016/j.envsoft.2015.10.025>
- [64] Paula A. Harrison, Robert W. Dunford, Ian P. Holman, and Mark Rounsevell. 2016. Climate change impact modelling needs to include cross-sectoral interactions. *Nature Climate Change* 6 (2016), 885–890. Issue 9. <https://doi.org/10.1038/nclimate3039>
- [65] Rex Hartson. 2003. Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour & Information Technology* 22, 5 (2003), 315–338. <https://doi.org/10.1080/01449290310001592587>
- [66] Peter A. Henrys and Susan G. Jarvis. 2019. Integration of ground survey and remote sensing derived data: Producing robust indicators of habitat extent and condition. *Ecology and Evolution* 9, 14 (2019), 8104–8112. <https://doi.org/10.1002/ece3.5376>
- [67] Morten Hertzum. 1999. Six Roles of Documents in Professionals’ Work. In *Proceedings of the Sixth European Conference on Computer Supported Cooperative Work* (Copenhagen, Denmark). Kluwer Academic Publishers, USA, 41–60. [https://doi.org/10.1007/978-94-011-4441-4\\_3](https://doi.org/10.1007/978-94-011-4441-4_3)
- [68] Michael Hind, Stephanie Houde, Jacquelyn Martino, Aleksandra Mojsilovic, David Piorowski, John Richards, and Kush R. Varshney. 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* (Honolulu, HI, USA) (CHI EA '20). Association for Computing Machinery, New York, NY, USA, 1–8. <https://doi.org/10.1145/3334480.3383051>
- [69] Michal R. Hoffman, Luis-Daniel Ibáñez, Huw Fryer, and Elena Simperl. 2018. Smart Papers: Dynamic Publications on the Blockchain. In *The Semantic Web, Aldo Gangemi, Roberto Navigli, Maria-Esther Vidal, Pascal Hitzler, Raphaël Troncy, Laura Hollink, Anna Tordai, and Mehwish Alam* (Eds.). Springer International Publishing, Cham, 304–318. [https://doi.org/10.1007/978-3-319-93417-4\\_20](https://doi.org/10.1007/978-3-319-93417-4_20)
- [70] Sarah Holland, Ahmed Hosny, Sarah Newman, Joshua Joseph, and Kasia Chmielinski. 2018. The Dataset Nutrition Label: A Framework To Drive Higher Data Quality Standards. arXiv:1805.03677
- [71] Michael Holloway, Graham Dean, Gordon Blair, Mike Brown, Pete Henrys, and John Watkins. 2020. Tackling the Challenges of 21st Open Science and Beyond: A Data Science Lab Approach. *Patterns* (2020). <https://doi.org/10.1016/j.patter.2020.100103>
- [72] Jennifer M. Holzer, Cristian Mihai Adamescu, Constantin Cazacu, Ricardo Diaz-Delgado, Jan Dick, Pablo F. Méndez, Luis Santamaría, and Daniel E. Orenstein. 2019. Evaluating transdisciplinary science to open research-implementation spaces in European social-ecological systems. *Biological Conservation* 238 (2019), 108228. <https://doi.org/10.1016/j.biocon.2019.108228>
- [73] Ian Hutchby. 2001. Technologies, Texts and Affordances. *Sociology* 35, 2 (2001), 441–456. <https://doi.org/10.1177/S0038038501000219>
- [74] Anna Sigridur Islind, Ulrika Lundh Snis, Tomas Lindroth, Johan Lundin, Katerina Cerna, and Gunnar Steineck. 2019. The Virtual Clinic: Two-Sided Affordances in Consultation Practice. *Comput. Supported Coop. Work* 28, 3–4 (June 2019), 435–468. <https://doi.org/10.1007/s10606-019-09350-3>
- [75] Corey Jackson, Kevin Crowston, Carsten Østerlund, and Mahboobeh Harandi. 2018. Folksonomies to Support Coordination and Coordination of Folksonomies. *Comput. Supported Coop. Work* 27, 3–6 (Dec. 2018), 647–678. <https://doi.org/10.1007/s10606-018-9327-z>
- [76] Udo Kannegiesser and John S Gero. 2012. A Process Framework of Affordances in Design. *Design Issues* 28, 1 (2012), 50–62. [https://doi.org/10.1162/DESI\\_a\\_00123](https://doi.org/10.1162/DESI_a_00123)
- [77] Victor Kaptelinin 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) (CHI '12). Association for Computing Machinery, New York, NY, USA, 967–976. <https://doi.org/10.1145/2207676.2208541>
- [78] Jakko Kemper and Daan Kolkman. 2019. Transparent to whom? No algorithmic accountability without a critical audience. *Information, Communication & Society* 22, 14 (2019), 2081–2096. <https://doi.org/10.1080/1369118X.2018.1477967>
- [79] Kimia Kiani, George Cui, Andrea Bunt, Joanna McGrenere, and Parmit K. Chilana. 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). Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3290605.3300570>

- [80] Iacovos Kirlappos, Martina Angela Sasse, and Nigel Harvey. 2012. Why Trust Seals Don't Work: A Study of User Perceptions and Behavior. In *Trust and Trustworthy Computing - 5th International Conference, TRUST 2012, Vienna, Austria, June 13-15, 2012. Proceedings (Lecture Notes in Computer Science, Vol. 7344)*, Stefan Katzenbeisser, Edgar R. Weippl, L. Jean Camp, Melanie Volkamer, Mike K. Reiter, and Xinwen Zhang (Eds.). Springer, 308–324. [https://doi.org/10.1007/978-3-642-30921-2\\_18](https://doi.org/10.1007/978-3-642-30921-2_18)
- [81] Bran Knowles. 2016. Emerging Trust Implications of Data-Rich Systems. *IEEE Pervasive Computing* 15, 4 (Oct. 2016), 76–84. <https://doi.org/10.1109/MPRV.2016.68>
- [82] Bran Knowles, Mike Harding, Lynne Blair, Nigel Davies, James Hannon, Mark Rouncefield, and John Walden. 2014. Trustworthy by Design. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work Social Computing* (Baltimore, Maryland, USA) (CSCW '14). Association for Computing Machinery, New York, NY, USA, 1060–1071. <https://doi.org/10.1145/2531602.2531699>
- [83] Bran Knowles and John T. Richards. 2021. The Sanction of Authority: Promoting Public Trust in AI. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAcCT '21)*. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3442188.3445890>
- [84] Bran Knowles, Mark Rouncefield, Mike Harding, Nigel Davies, Lynne Blair, James Hannon, John Walden, and Ding Wang. 2015. Models and Patterns of Trust. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work Social Computing* (Vancouver, BC, Canada) (CSCW '15). Association for Computing Machinery, New York, NY, USA, 328–338. <https://doi.org/10.1145/2675133.2675154>
- [85] Guiou Kobayashi, Mariana Claudia Broens, Maria Eunice Quilici Gonzalez, and Jose Artur Quilici-Gonzalez. 2015. The internet of things and its impact on social relationships involving mutual trust. In *2015 IEEE International Symposium on Technology and Society (ISTAS)*, 1–6. <https://doi.org/10.1109/ISTAS.2015.7439416>
- [86] Károly Zoltán Kovács, Drew Hemment, Mel Woods, Naomi K. van der Velden, Angelika Xaver, Rianne H. Gi Eesen, Victoria J. Burton, Natalie L. Garrett, Luca Zappa, Deborah Long, Endre Dobos, and Rastislav Skalsky. 2019. Citizen observatory based soil moisture monitoring – The GROW example. *Hungarian Geographical Bulletin* 68, 2 (1 Jan. 2019), 119–139. <https://doi.org/10.15201/hungeobull.68.2.2>
- [87] Frens Kroeger. 2017. FACEWORK: creating trust in systems, institutions, and organisations. *Cambridge Journal of Economics* 41, 2 (1 3 2017), 487–514. <https://doi.org/10.1093/cje/bew038>
- [88] Frens Kroeger. 2019. Unlocking the treasure trove: How can Luhmann's theory of trust enrich trust research? *Journal of Trust Research* 9, 1 (2019), 110–124. <https://doi.org/10.1080/21515581.2018.1552592>
- [89] Anna Krzywoszynska. 2019. Making knowledge and meaning in communities of practice: What role may science play? The case of sustainable soil management in England. *Soil Use and Management* 35, 1 (2019), 160–168. <https://doi.org/10.1111/sum.12487>
- [90] Dong Joon Lee and Besiki Stvilja. 2017. Practices of research data curation in institutional repositories: A qualitative view from repository staff. *PLOS ONE* 12, 3 (03 2017), 1–44. <https://doi.org/10.1371/journal.pone.0173987>
- [91] John D. Lee and Katrina A. See. 2004. Trust in Automation: Designing for Appropriate Reliance. *Human Factors* 46, 1 (2004), 50–80. [https://doi.org/10.1518/hfes.46.1.50\\_30392](https://doi.org/10.1518/hfes.46.1.50_30392)
- [92] Min Kyung Lee. 2018. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society* 5, 1 (2018), 2053951718756684. <https://doi.org/10.1177/2053951718756684>
- [93] Paul M. Leonardi. 2011. When Flexible Routines Meet Flexible Technologies: Affordance, Constraint, and the Imbrication of Human and Material Agencies. *MIS Quarterly* 35, 1 (2011), 147–167. <http://www.jstor.org/stable/23043493>
- [94] J. David Lewis and Andrew Weigert. 1985. Trust as a Social Reality. *Social Forces* 63, 4 (06 1985), 967–985. <https://doi.org/10.1093/sf/63.4.967>
- [95] Brian Y. Lim and Anind K. Dey. 2013. Evaluating Intelligibility Usage and Usefulness in a Context-Aware Application. In *Human-Computer Interaction. Towards Intelligent and Implicit Interaction - 15th International Conference, HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part V (Lecture Notes in Computer Science, Vol. 8008)*, Masaaki Kurosu (Ed.). Springer, 92–101. [https://doi.org/10.1007/978-3-642-39342-6\\_11](https://doi.org/10.1007/978-3-642-39342-6_11)
- [96] Philip Lord, Alison Macdonald, Liz Lyon, and David Giarretta. 2004. From Data Deluge to Data Curation. In *Proceedings of the UK E-Science All Hands Meeting*, 371–375.
- [97] Chloe Lucas, Peat Leith, and Aidan Davison. 2015. How climate change research undermines trust in everyday life: a review. *WIREs Climate Change* 6, 1 (2015), 79–91. <https://doi.org/10.1002/wcc.320>
- [98] Niklas Luhmann. 2017. *Trust and Power*. Polity Press, Cambridge.
- [99] Tatiana Lukoianova and Victoria Rubin. 2014. Veracity Roadmap: Is Big Data Objective, Truthful and Credible? *Advances in Classification Research Online* 24, 1 (2014), 4–15. <https://doi.org/10.7152/acro.v24i1.14671>
- [100] John Lyle and Andrew Martin. 2010. Trusted Computing and Provenance: Better Together. In *Proceedings of the 2nd Conference on Theory and Practice of Provenance* (San Jose, California) (TAPP'10). USENIX Association, USA. <http://www.cs.ox.ac.uk/files/5278/tapp10-cameraready.pdf>
- [101] Stuart Macdonald and Luis Martinez-Urbe. 2010. Collaboration to Data Curation: Harnessing Institutional Expertise. *New Review of Academic Librarianship* 16, sup1 (2010), 4–16. <https://doi.org/10.1080/13614533.2010.505823>
- [102] Stephen Marsh, Tosan Atele-Williams, Anirban Basu, Natasha Dwyer, Peter R. Lewis, Hector Miller-Bakewell, and Jeremy Pitt. 2020. Thinking about Trust: People, Process, and Place. *Patterns* 1, 3 (2020). <https://doi.org/10.1016/j.patter.2020.100039>
- [103] Katherine Mattor, Michele Betsill, Ch'aska Huayhuaca, Heidi Huber-Stearns, Theresa Jedd, Faith Sternlieb, Patrick Bixler, Matthew Luizza, and Antony Cheng. 2014. Transdisciplinary research on environmental governance: A view from the inside. *Environmental Science Policy* 42 (10 2014), 90–100. <https://doi.org/10.1016/j.envsci.2014.06.002>
- [104] Roger C. Mayer, James H. Davis, and F. David Schoorman. 1995. An Integrative Model of Organizational Trust. *The Academy of Management Review* 20, 3 (1995), 709–734. <http://www.jstor.org/stable/258792>
- [105] Ramia Mazé and Johan Redström. 2005. Form and the computational object. *Digital Creativity* 16, 1 (2005), 7–18. <https://doi.org/10.1080/14626260500147736>
- [106] Joanna McGrenere and Wayne Ho. 2000. Affordances: Clarifying and Evolving a Concept. In *Proceedings of the Graphics Interface 2000 Conference, May 15-17, 2000, Montréal, Québec, Canada*, 179–186. <https://doi.org/10.20380/GI2000.24>
- [107] D. Harrison McKnight, Vivek Choudhury, and Charles Kacmar. 2002. Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research* 13, 3 (2002), 334–359. <https://doi.org/10.1287/isre.13.3.334.81>
- [108] Carsten Meyer, Patrick Weigelt, and Holger Kreft. 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecology Letters* 19, 8 (2016), 992–1006. <https://doi.org/10.1111/ele.12624>
- [109] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) (FAT'19). Association for Computing Machinery, New York, NY, USA, 220–229. <https://doi.org/10.1145/3287560.3287596>
- [110] Brent Mittelstadt, Chris Russell, and Sandra Wachter. 2019. Explaining Explanations in AI. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) (FAT'19). Association for Computing Machinery, New York, NY, USA, 279–288. <https://doi.org/10.1145/3287560.3287574>
- [111] Guido Möllering. 2001. The Nature of Trust: From Georg Simmel to a Theory of Expectation, Interpretation and Suspension. *Sociology* 35, 2 (2001), 403–420. <https://doi.org/10.1177/S0038038501000190>
- [112] Donald A. Norman. 1999. Affordance, Conventions, and Design. *Interactions* 6, 3 (1999), 38–43. <https://doi.org/10.1145/301153.301168>
- [113] Donald A. Norman. 2002. *The Design of Everyday Things*. Basic Books, New York.
- [114] Vatsala Nundloll, Barry Porter, Gordon Blair, Bridget Emmett, Davey Jones, Dave Chadwick, Ben Winterbourn, Philip Beattie, Graham Dean, Rory Shaw, Wayne Shelley, Mike Brown, and Izhar Ullah. 2019. The Design and Deployment of an End-To-End IoT Infrastructure for the Natural Environment. *Future Internet* 11, 6 (2019). <https://doi.org/10.3390/fi11060129>
- [115] Kieron O'Hara. 2019. *Data Trusts: Ethics, Architecture and Governance for Trustworthy Data Stewardship*. Project Report 10.5258/SOTON/WSI-WP001. <https://doi.org/doi:10.5258/SOTON/WSI-WP001>
- [116] Martin Oliver. 2005. The Problem with Affordance. *E-Learning and Digital Media* 2, 4 (2005), 402–413. <https://doi.org/10.2304/elea.2005.2.4.402>
- [117] Onora O'Neill. 2002. *A Question of Trust: The BBC Reith Lectures*. Cambridge University Press, Cambridge.
- [118] Onora O'Neill. 2018. Linking Trust to Trustworthiness. *International Journal of Philosophical Studies* 26, 2 (2018), 293–300. <https://doi.org/10.1080/09672559.2018.1454637>
- [119] Naomi Oreskes. 2019. *Why Trust Science?* Princeton University Press, Princeton, NJ. <https://doi.org/10.2307/j.ctvfjczxx>
- [120] Wanda J. Orlikowski. 2000. Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations. *Organization Science* 11, 4 (2000), 404–428. <https://doi.org/10.1287/orsc.11.4.404.14600>
- [121] Samir Passi and Steven J. Jackson. 2018. Trust in Data Science: Collaboration, Translation, and Accountability in Corporate Data Science Projects. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (2018). <https://doi.org/10.1145/3274405>
- [122] Oliver Pescott, Thomas Humphrey, and Kevin Walker. 2018. A short guide to using British and Irish plant occurrence data for research. (07 2018). <https://doi.org/10.13140/RG.2.2.33746.86720>
- [123] Wolter Pieters. 2011. Explanation and Trust: What to Tell the User in Security and AI? *Ethics and Inf. Technol.* 13, 1 (March 2011), 53–64. <https://doi.org/10.1007/s10676-010-9253-3>
- [124] Theodore M. Porter. 1995. *Trust in numbers the pursuit of objectivity in science and public life*. Princeton University Press, Princeton, NJ.

- [125] Forough Poursabzi-Sangdeh, Daniel G. Goldstein, Jake M. Hofman, Jennifer Wortman Vaughan, and Hanna M. Wallach. 2018. Manipulating and Measuring Model Interpretability. (2018). arXiv:1802.07810
- [126] Emilee Rader, Kelley Cotter, and Janghee Cho. 2018. Explanations as Mechanisms for Supporting Algorithmic Transparency. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3173677>
- [127] Madhu Reddy, Paul Dourish, and Wanda Pratt. 2001. Coordinating Heterogeneous Work: Information and Representation in Medical Care. *European Conference on Computer Supported Cooperative Work (ECSCW'01)* (2001), 239–258. [https://doi.org/10.1007/0-306-48019-0\\_13](https://doi.org/10.1007/0-306-48019-0_13)
- [128] Derek Reilly and Bonnie MacKay. 2013. Annotating Ecology: Looking to Biological Fieldwork for Mobile Spatial Annotation Workflows. In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Munich, Germany) (MobileHCI '13). Association for Computing Machinery, New York, NY, USA, 35–44. <https://doi.org/10.1145/2493190.2493245>
- [129] John Richards, David Piorkowski, Michael Hind, Stephanie Houde, and Aleksandra Mojsilović. 2020. A Methodology for Creating AI FactSheets. arXiv:2006.13796
- [130] Gernot Rieder and Judith Simon. 2016. Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data. *Big Data & Society* 3, 1 (2016), 2053951716649398. <https://doi.org/10.1177/2053951716649398>
- [131] Jens Riegelsberger, M. Angela Sasse, and John D. McCarthy. 2005. The Mechanics of Trust: A Framework for Research and Design. *Int. J. Hum.-Comput. Stud.* 62, 3 (March 2005), 381–422. <https://doi.org/10.1016/j.ijhcs.2005.01.001>
- [132] Jeremy Rose and Rens Scheepers. 2001. Structuration Theory and Information System Development - Frameworks for Practice. In *9th European Conference on Information Systems (ECIS 2001)*. Association for Information Systems, 217–231. <https://aisel.laisnet.org/ecis2001/80>
- [133] Denise M. Rousseau, Sim B. Sitkin, Ronald S. Burt, and Colin Camerer. 1998. Not So Different After All: A Cross-Discipline View Of Trust. *Academy of Management Review* 23, 3 (1998), 393–404. <https://doi.org/10.5465/amr.1998.926617>
- [134] Johannes Sanger, Christian Richthammer, Sabri Hassan, and Gunther Pernul. 2014. Trust and Big Data: A Roadmap for Research. In *Proceedings of the 2014 International Semiconductor Laser Conference (ISLC '14)*. IEEE Computer Society, USA, 278–282. <https://doi.org/10.1109/DEXA.2014.63>
- [135] Andrew Schrock. 2015. Communicative Affordances of Mobile Media: Portability, Availability, Locatability, and Multimediality. *International Journal of Communication* 9 (2015), 1229–1246. <https://ijoc.org/index.php/ijoc/article/view/3288>
- [136] Anna Sexton, Elizabeth Shepherd, Oliver Duke-Williams, and Alexandra Eveleigh. 2005. A Balance of Trust in the Use of Government Administrative Data. *Archival Science* 17, 4 (2005), 305 – 330. <https://doi.org/10.1007/s10502-017-9281-4>
- [137] Donghee Shin, Bu Zhong, and Frank A. Biocca. 2020. Beyond user experience: What constitutes algorithmic experiences? *International Journal of Information Management* 52 (2020). <https://doi.org/10.1016/j.ijinfomgt.2019.102061>
- [138] Clarisse Sieckenius de Souza, Raquel Oliveira Prates, and Tom Carey. 2000. Missing and declining affordances: are these appropriate concepts? *Journal of the Brazilian Computer Society* 7 (00 2000), 26 – 34. [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0104-65002000000200004&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-65002000000200004&nrm=iso)
- [139] Raphael Spiekermann, Ben Jolly, Alexander Herzig, Tom Burleigh, and David Medyckyj-Scott. 2019. Implementations of fine-grained automated data provenance to support transparent environmental modelling. *Environmental Modelling Software* 118 (2019), 134–145. <https://doi.org/10.1016/j.envsoft.2019.04.009>
- [140] Alex Stevens. 2011. Telling Policy Stories: An Ethnographic Study of the Use of Evidence in Policy-making in the UK. *Journal of Social Policy* 40, 2 (2011), 237–255. <https://doi.org/10.1017/S0047279410000723>
- [141] Lauren Thornton, Victoria Neumann, Gordon Blair, Nigel Davies, and John Watkins. 2019. Trusted Brokers?: Identifying the Challenges Facing Data Centres. Data for Policy. <https://doi.org/10.5281/zenodo.2798468>
- [142] Ehsan Toreini, Mhairi Aitken, Kovila Coopamootoo, Karen Elliott, Carlos Gonzalez Zelaya, and Aad van Moorsel. 2020. The Relationship between Trust in AI and Trustworthy Machine Learning Technologies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) (FAT\* '20). Association for Computing Machinery, New York, NY, USA, 272–283. <https://doi.org/10.1145/3351095.3372834>
- [143] Gerard Torenvliet. 2003. We can't afford it!: the devaluation of a usability term. *Interactions* 10, 4 (2003), 12–17. <https://doi.org/10.1145/838830.838857>
- [144] Ross Towe, Graham Dean, Liz Edwards, Vatsala Nundloll, Gordon Blair, Rob Lamb, Barry Hankin, and Susan Manson. 2020. Rethinking data-driven decision support in flood risk management for a big data age. *Journal of Flood Risk Management* (2020). <https://doi.org/10.1111/jfr3.12652>
- [145] Dhaval Vyas, Cristina M. Chisalita, and Alan Dix. 2017. Organizational Affordances: A Structuration Theory Approach to Affordances. *Interacting with Computers* 29, 2 (2017), 117–131. <https://doi.org/10.1093/iwc/iww008>
- [146] Dhaval Vyas, Cristina M. Chisalita, and Gerrit C. van der Veer. 2006. Affordance in Interaction. In *Proceedings of the 13th European Conference on Cognitive Ergonomics: Trust and Control in Complex Socio-Technical Systems* (Zurich, Switzerland) (ECCE '06). Association for Computing Machinery, New York, NY, USA, 92–99. <https://doi.org/10.1145/1274892.1274907>
- [147] Dhaval Vyas and Gerrit C. van der Veer. 2006. Experience as Meaning: Some Underlying Concepts and Implications for Design. In *Proceedings of the 13th European Conference on Cognitive Ergonomics: Trust and Control in Complex Socio-Technical Systems* (Zurich, Switzerland) (ECCE '06). Association for Computing Machinery, New York, NY, USA, 81–91. <https://doi.org/10.1145/1274892.1274906>
- [148] Jillian Wallis, Christine Borgman, Matthew Mayernik, Alberto Pepe, R.Anitha Nithya, and Mark Hansen. 2007. Know Thy Sensor: Trust, Data Quality, and Data Integrity in Scientific Digital Libraries, Vol. 4675. 380–391. [https://doi.org/10.1007/978-3-540-74851-9\\_32](https://doi.org/10.1007/978-3-540-74851-9_32)
- [149] Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y. Lim. 2019. Designing Theory-Driven User-Centric Explainable AI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3290605.3300831>
- [150] Nicholas M. Weber, Carole L. Palmer, and Tiffany C. Chao. 2012. Current Trends and Future Directions in Data Curation Research and Education. *Journal of Web Librarianship* 6, 4 (2012), 305–320. <https://doi.org/10.1080/19322909.2012.730358>
- [151] Maranke Wieringa. 2020. What to Account for When Accounting for Algorithms: A Systematic Literature Review on Algorithmic Accountability. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) (FAT\* '20). Association for Computing Machinery, New York, NY, USA, 1–18. <https://doi.org/10.1145/3351095.3372833>
- [152] Mark Wilkinson, Eleanor Mackay, Paul F. Quinn, Marc Stutter, Keith Beven, Christopher J.A. Macleod, M. G. Macklin, Yehia Elkhatib, Barbara J. Percy, Claudia Vitolo, and Philip Haygarth. 2015. A cloud based tool for knowledge exchange on local scale flood risk. *Journal of Environmental Management* 161 (2015), 38–50. <https://doi.org/10.1016/j.jenvman.2015.06.009>
- [153] Emily Winter, Maria Angela Ferrario, and Gordon Blair. 2020. Agility in Transdisciplinary Research: Lessons Learnt from a Research Sprint on Digital Technologies and Flood Risk Management. *International Conferences of ICT for Sustainability (ICT4S)* (2020). [https://eprints.lancs.ac.uk/id/eprint/144673/1/ICT4S\\_FINAL\\_21\\_MAY.pdf](https://eprints.lancs.ac.uk/id/eprint/144673/1/ICT4S_FINAL_21_MAY.pdf)
- [154] Ming Yin, Jennifer Wortman Vaughan, and Hanna Wallach. 2019. Understanding the Effect of Accuracy on Trust in Machine Learning Models. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3290605.3300509>
- [155] Meg Young, Luke Rodriguez, Emily Keller, Feiyang Sun, Boyang Sa, Jan Whittington, and Bill Howe. 2019. Beyond Open vs. Closed: Balancing Individual Privacy and Public Accountability in Data Sharing. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) (FAT\* '19). Association for Computing Machinery, New York, NY, USA, 191–200. <https://doi.org/10.1145/3287560.3287577>
- [156] Jun Zhao, Oscar Corcho, Paolo Missier, Khalid Belhajjame, David Newmann, David de Roure, and Carole A. Goble. 2011. eScience. Springer Berlin Heidelberg, Berlin, Heidelberg, 701–736. [https://doi.org/10.1007/978-3-540-92913-0\\_17](https://doi.org/10.1007/978-3-540-92913-0_17)
- [157] Jun Zhao, Jose Manuel Gomez-Perez, Khalid Belhajjame, Graham Klyne, Esteban Garcia-Cuesta, Alex Garrido, Kristina Hettne, Marco Roos, David De Roure, and Carole A. Goble. 2012. Why workflows break – Understanding and combating decay in Taverna workflows. In *2012 IEEE 8th International Conference on e-Science*. 1–9. <https://doi.org/10.1109/eScience.2012.6404482>
- [158] Anneke Zuiderwijk. 2017. Analysing Open Data in Virtual Research Environments: New Collaboration Opportunities to Improve Policy Making. *Int. J. Electron. Gov. Res.* 13, 4 (Oct. 2017), 76–92. <https://doi.org/10.4018/IJEGR.2017100105>