Archives & Records, 2013
Vol. 34, No. 2, 214–234, http://dx.doi.org/10.1080/23257962.2013.825207

Routledge
Taylor & Francis Group

Trust and context in cyberspace

Geoffrey Yeo*

Department of Information Studies, University College London, Gower Street, London WC1E 6BT, UK

Every day we place trust or reliance on other people and on inanimate objects, but trust may be diminished in the world of information resources and technology. We are often told that information needs higher standards of verification in digital realms than in the paper world. Similarly, when we encounter digital records and archives we may be uncertain how far we can trust them. In the past, trust in records was said to be reinforced by trust in archivists and archival institutions. However, trust in professional experts and institutions is waning; notions of expert objectivity are increasingly challenged. This paper explores an idea proposed by David Weinberger, that 'transparency is the new objectivity'. Where records are concerned, documentation of provenance and context forms a basis for enhancing their transparency and thus for evaluating their trustworthiness. Many commentators have expressed anxiety that, in digital environments where resources are reused and remixed at will, records may become decontextualized. But in computer science questions are now being asked about how *data* can be trusted and verified, and knowledge of their provenance is increasingly seen as a foundation for enabling trust. Many computer scientists argue that, while data should be reusable, each piece of data should carry evidence of its history and original contexts to help those who encounter it to judge its trustworthiness. Some researchers have set out to develop systems to capture and preserve information about data provenance. In the longer term, this research may help archivists meet the challenges of gathering and maintaining contextual information in the world of digital record-keeping. Methods of automatically harvesting certain kinds of contextual information are under investigation; automated solutions are likely to expedite what are currently time-consuming manual processes. However, merely being presented with information about provenance is not enough. Insofar as individuals or institutions supply us with that information, we have to decide how far we trust what those people or institutions tell us. There is still a place for expert voices, but experts cannot be seen as infallible providers of objective information.

Keywords: trust; context; provenance; transparency; digital records

Introduction

In the late summer of 2012, Professor Luciana Duranti invited me to give the opening public address at the Peter Wall Exploratory Workshop on Trust and Conflicting Rights in the Digital Environment.[1] The address was to be delivered to a mixed audience, including records managers, archivists and people who were not archival professionals. The topic she asked me to speak on was 'Trust in Cyberspace'. Of course, this is potentially a huge topic, and I did not attempt to explore all its many ramifications; instead, I chose to focus on particular themes of concern to keepers and users of records. This paper is the published version of my address.[2] I have tidied it up a little for publication, but I have not attempted to adjust the emphasis or extend the scope of the paper to provide a fuller account of the

*Email: g.yeo@ucl.ac.uk

complexities of trust.[3] However, I have amended the title of the paper to 'Trust and Context in Cyberspace', to align it more closely to the themes I selected. This paper first discusses aspects of trust in everyday life and in interactions with information resources and technology, and considers the role played by understandings of context and provenance in establishing trust in records and archives. It then examines the recent growth of interest in transparency, trust, provenance and context among computer scientists, and explores ways in which trust and provenance research in computer science may be able to help archivists meet the challenges of capturing contextual information in the age of digital records.

Trust

Trust, the philosopher John Locke affirmed in 1663, is 'the bond of society',[4] and we may well be inclined to agree with him. On a daily basis, we find that we trust others. If we get into a taxi we trust the driver to take us where we have told him we want to go; we trust that he knows how to drive the taxi safely and that other drivers can also be relied on to drive in a careful manner so that he (and we) will not be involved in an accident. If a police officer indicates that the driver must stop, or must follow a diversion, we trust that she has a good reason for doing so. Social life would scarcely function if we did not feel able to exercise such trust.

We also trust, or at least rely on, inanimate objects. We trust that the taxi itself, its safety belts and tyres and headlights will all operate as they should and perform the tasks we expect of them. In fact, we rely on the effectiveness of vast numbers of artefacts, from alarm clocks to thermometers to bridges to missile early-warning systems. We take a similar attitude to the natural world. We trust that the fruit we eat will not make us ill, that the tree we climb will bear our weight and that the rocks above our path will not fall on us. We know that there is always a possibility, however small, that the clock will not wake us, that the taxi will break down or that the rocks will tumble onto our heads; but in practice this knowledge does not dissuade us from sleeping or making our journey. We are aware that banknotes can be forged, but we do not let this awareness deter us from accepting them. Generally, we do not even consider the possibility that our trust might be misplaced; we go ahead and use the taxi or the alarm clock or the €10 note without giving the matter any further thought. It is possible to argue that in these situations trust is a matter of risk assessment: if the risk is low enough, we will decide to trust the object or artefact concerned.[5] But usually we are not conscious that we assess a risk; we are simply getting on with our lives.

Records are another kind of inanimate object that we trust. A record is a representation of a past activity that we can use to obtain evidence, or reinforce memory, of what happened in the past. In general, as far as trust is concerned, our attitude to records is not very different from our attitude to taxis, alarm clocks or €10 notes. We use them when we need them; we are aware that records can sometimes be fraudulent or unreliable, but in practice we do not always take the trouble to verify the records we use. Sometimes, however, we seek to reassure ourselves that particular records can be trusted; later in this paper I will examine some approaches that we might follow when such reassurance is required.

In addition, we trust, or have trusted in the past, our institutions and our governments.[6] Universities, churches, parliaments and courts of law all to some degree depend on public trust for their continuing legitimacy and effectiveness. But there is evidence that this kind of trust may now be in decline. Researchers have found substantial falls in levels of trust

towards governmental institutions in recent decades in many western countries, including Sweden, the UK and the USA.[7] Since the start of the latest financial crisis, journalists have frequently noted that banking corporations have lost much of the trust that their clients once gave them.[8] In the 1990s, the work of the political scientist Francis Fukuyama appeared to suggest that trust which depends on ethical habits and reciprocal moral obligations is undermined in modernist societies where economic rationality is considered the sole or main arbiter of social organization.[9] Other writers have linked the decline in trust to marketization and the growing inequalities to which it gives rise, and have pointed to surveys showing that levels of trust seem to be highest in countries with low levels of social inequality.[10]

The financial crisis has undoubtedly stimulated new debate about trust. Many commentators have suggested that increased regulation is needed to rebuild trust in the banking industry, but others have argued that the resort to more and tighter regulation is symptomatic of a society where trust is lacking, a society based on a culture of suspicion. From this point of view, the introduction of rules and audits to promote trust merely serve to erode it.[11] Of course, bankers are not the only targets of the audit culture. Indeed, before the financial crisis, many neoliberal policy-makers had argued that banks and bankers should be largely exempt from regulation. An easier and more suitable target for regulatory control was the professional: the teacher, the doctor, the social worker or the academic. It is probably not fortuitous that the movement towards increased regulation of these expert professions has largely coincided with the rise of postmodern antipathy to expertise, and with a growing emphasis on multiplicity of perspectives and democratization of knowledge. Just as the days are gone when the banker's word was trusted as his bond, so also the days are gone when the word of the professional was invariably trusted as authoritative and objective.

Cyberspace

A further area where trust is often said to be at risk is the territory of cyberspace. Of course, 'cyberspace' is one of those gibberish words that make many of us sigh when we hear it; I will use it merely as a convenient shorthand for the world of computer technology, and especially those parts of it associated with the Internet. In this world, trust may seem to be diminished because communication is depersonalized and the identity of actors is hidden behind websites or other digital proxies.[12] As the information scientist Clifford Lynch observed, 'many people feel that in an environment characterized by pervasive deceit, it … [is] necessary to provide … proof for claims … that would usually be taken at face value in the physical world'.[13]

We must not exaggerate the novelty of deceit in the digital domain. We all know about the fraudulent e-mails that try to solicit our assistance in retrieving a hidden fortune or securing an inheritance, but we may be less familiar with the equally fictitious appeals that were common in nineteenth-century paper mail, when writers of letters claimed that a man wrongfully imprisoned in a foreign country needed a sum of money to obtain his freedom and would pay a generous reward or reveal the site of buried treasure on his release.[14] Untrustworthy communications are not a novelty. What is new is not the nature of the deception, but the way in which Internet technology allows fraudsters to operate with relative ease and to disseminate their scams widely at minimal cost to themselves.

It is often thought that trust issues in cyberspace mostly relate to social computing or e-commerce (how far can I trust the person I am chatting with or buying from?), but there are also concerns about software applications (can I trust software not to behave maliciously?)

and about the data and information that are purveyed to us in the digital realm. Suspicions about the latter can arise partly because of the absence of many of the cues we rely on in assessing non-digital information, partly because of the increased scope for deception and partly because of the ease with which it is now possible to publish information that seems to be unsubstantiated, mistaken or improbable.[15] To quote Lynch again, we often have 'a sense that digital information needs to be held to a higher standard for authenticity and integrity' than information in the paper world.[16] Tim Berners-Lee, the inventor of the World Wide Web, seems to have shared this view; his vision of an ideal web browser included what he called an 'Oh yeah?' button, a button that when invoked would give reasons why a web page or service should be believed.[17]

Of course, further difficulties arise from increasing scepticism about the notion that pieces of information can be definitively identified as true or false, as deserving or undeserving of belief. In 2002, Statistics Canada (the statistical agency of the Canadian government) asserted that the quality of the information it provides is 'critical to the Agency's reputation as an independent, objective source of trustworthy information', a claim that can still be found on its website today.[18] Ultimately, discussion of statements like this is likely to oblige us to take positions on controversial questions of ontology and epistemology; but what seems certain is that we can no longer ignore the many critics who insist that objectivity is discredited and cannot meaningfully be attributed either to a provider of information or to information itself. In the words of David Weinberger, a well-known commentator on developments in information technology, 'if you don't think objectivity is possible, then you think that the claim of objectivity is actually hiding the biases that inevitably are there'.[19]

We meet these same concerns, or very similar concerns, when we consider records and archives in cyberspace. Many of us, and many of our employers, now create, store and access records in Internet environments, often using cloud computing infrastructures established by commercial suppliers whose interests may not coincide with our own. While we usually welcome the convenience these environments offer, when we encounter records – or objects that purport to be records – in the digital realm we may be unsure how far we can trust what is being purveyed to us.[20]

We may also be uncertain how far traditional methods of verifying trustworthiness can be applied in these new domains. Because paper and parchment were for so long the media of choice for the creation and transmission of records in western societies, many methods have been developed for evaluating or enhancing the trustworthiness of records in these media.[21] Some of these methods can be applied to digital records and others have near-equivalents in the digital world; much of the functionality of seals and signatures, for example, can be replicated using cryptographic techniques. Nevertheless, we are likely to feel that digital signatures neither fulfil all the functions of written signatures nor have the same level of acceptance in the world at large; to the man or woman in the street, a written signature is a familiar device but a digital signature is an obscure mechanism understood and used only by technologists.[22] Much the same can be said of diplomatic or forensic examinations of records to assess their trustworthiness; in both paper and digital environments, most of us can make some kind of judgement whether a letter or e-mail we have received looks suspicious, but advanced diplomatic and forensic techniques are largely inaccessible to the non-specialist.

In the past, trust in archival records was said to be reinforced by trust in archivists and in the institutions where archives were kept. Archival institutions were relied on to preserve records in a way that inspired trust, and individuals who lacked the knowledge needed to evaluate a record could call on professionals to perform the necessary authentication.[23] Archivists and archival institutions were seen as neutral and objective

third parties that could be trusted to protect records and not tamper with them.[24] However, not only is trust in professional experts and institutions now waning, but we also face issues of disintermediation: online users cannot interact with archivists or sense the physical institution in the way that traditional users could. As we have seen, notions of expert objectivity are also increasingly challenged, and few of us are now confident that we can invariably cease questioning when we reach a credentialed professional who gives us an ostensibly objective answer.

Perhaps, then, we need new ways of dealing with these issues. Or perhaps we need to reinvent old ways to make them fit for the digital age. To pursue this line of enquiry, I want to build on an idea that has been put forward by David Weinberger. In a blog post in 2009, Weinberger claimed that 'transparency is the new objectivity'. More specifically, he argued that the role in 'the ecology of knowledge' once served by the unattainable goal of objectivity is now served by transparency. Transparency, he said, allows us to see how a resource was formed. It prospers in a linked medium such as the Internet because links let us see the connections between a resource and the ideas and values that informed it, and this in turn gives grounds for trust; it gives us reasons to have confidence in Internet resources in the way that claims about objectivity once did for paper materials. 'What we used to believe because we thought [an] author was objective', Weinberger affirmed, 'we now believe because we can see … the sources that brought her to that position'.[25]

So what can help to give *records* transparency in the world of the Internet? How can users who encounter a record see what shaped it and how it was formed? To help us move towards some answers to these questions, I will focus on just one set of ideas that already have a long history in archival science, and will explore how an approach that is familiar to archivists can support transparency and hence trust in cyberspace. I will also examine how recent initiatives in computer science can assist archivists in responding to some of the challenges they face in the world of digital record-keeping.

Provenance and context

The ideas I want to explore are those that relate to provenance. Broadly speaking, the principle of provenance requires records and archives to be managed in a way that secures and preserves knowledge of their origins and the circumstances of their creation. Among archivists, its methodological interpretation varies, but its theoretical validity is almost universally accepted. One writer has described it as 'the foundation on which modern archival science rests'.[26] While information about provenance is needed for practical management, archival literature tends to emphasize its role in aiding *comprehension* of archives, on the grounds that their content is rarely explicable without at least some knowledge of their provenance. Some writers have also contended that adherence to the principle of provenance underpins what Hilary Jenkinson called the 'moral defence' of archives, or that it safeguards their authenticity or integrity.[27]

Use of the term 'provenance' in connection with records and archives dates from the nineteenth century, but more recently it has been joined by another term that has become closely associated with it; the term 'context' is now as widely used in archival discourse as it is in literary studies and many other fields of scholarship. The precise relationships between the two terms are matters of debate. Some archivists tend to see provenance and context as more or less synonymous, while others claim that they are related but distinct. Some argue in favour of extending notions of provenance beyond the immediate origins of records to embrace their societal framing and the history of their custody and use,[28] while others insist that provenance is a more limited term and that context is a richer or subtler

notion referring to a wider range of phenomena that surround records or within which records are embedded.[29] Regardless of these disagreements, however, all parties agree that provenance and context are central concepts in archival discourse.

Research by Michael Roper and Margaret Procter has shown how, at the Public Record Office in London in the early 1900s, the principle of presenting records in accordance with their provenance or context overcame the previously prevailing view that only their information content was of value.[30] In the years immediately before the First World War, members of the Public Record Office's staff, perhaps influenced by the manual of Muller, Feith and Fruin, which had been published in 1898,[31] began to argue that researchers would wish to study records in their setting as products of administrative machinery and not merely as evidence of isolated facts. Many of today's researchers may be less sympathetic than their predecessors to the customs and practices of officialdom, but even researchers who want to read archives 'against the grain' may feel a need to understand the 'grain' they wish to counteract, in terms of the contextual milieus in which documents were produced.

But a century after the principle of provenance won acceptance at the Public Record Office, some critics have begun to express anxiety that the digital era may presage a return to a world of isolated facts used out of context. In the paper world, archivists developed ways of protecting the provenance of archives through methods of physical arrangement and written description, but many commentators have suggested that the protection these methods offered is at risk in digital environments where resources can be reused and remixed at will. In their recent book on relations between archives and history, Francis Blouin and William Rosenberg argued that when archival materials are 'structured apart from the contexts of their creation' they may be deprived 'of important kinds of historical meaning.'[32] Digital archivist Christopher Lee made a similar point when he discussed the reuse and aggregation of content from blogs. He borrowed an expression from Microsoft researcher Jonathan Grudin and spoke of the original conversation being 'desituated'; when RSS feeds, for example, deliver 'micro-content' from a blog, they strip out 'much of the contextual information that is so important'.[33]

Outside the archival field, the Internet governance specialist Viktor Mayer-Schönberger has also reminded us that 'circumstantial information … is largely missing from an information snippet directly accessed through digital retrieval'. Information, he affirmed, becomes de- and re-contextualized 'because pieces of information are retrieved without their accompanying contexts and presented in a new context of search results'. They are, or can be, 'severed from their original context, without a way to trace them back'.[34]

More than a decade ago, the sociologists Marc Berg and Els Goorman noted similar trends in the world of health care, where the shift to electronic patient records has been widely seen to betoken the liberation of information previously locked up in traditional paper-based records; electronic records of patient care are perceived as a source of information that can be extracted and reused for a range of secondary purposes including administration, financial management, service planning and medical research. Promoters of such reuse depict information as autonomous building blocks that can travel independently and be redeployed at will. Berg and Goorman argued that, while information reuse can be highly valuable, it is also dangerous, because information is 'always entangled with the context of its production'. They also noted that viewing pieces of information from health records as isolated 'atoms' overlooks their mutual elaboration. Rather than 'a heap of facts', the components of a patient's record are the 'bits and pieces of an emerging story'.[35] As archivists have long known, records (and parts of records)

have complex webs of interrelationships, which inform our judgements and understandings of the records; these judgements and understandings are at risk when context is ignored.[36]

But lately, Mayer-Schönberger remarked, 'software engineers around the world have been rushing to amend our digital tools to enable a modicum of (re)contextualization'.[37] He was right; archivists and sociologists are not alone in their concern that context should not be lost in the realm of cyberspace. Within the last few years, many computer scientists have discovered the word 'provenance', and much of their work on contextualization, or recontextualization, is explicitly concerned with what researchers in that field call the provenance of *data*. In the world of e-science and big data, where inconceivably huge quantities of scientific data can be continuously mashed up and endlessly remixed, questions are being asked about how these data can be trusted and verified, and knowledge of their provenance is increasingly seen as the key, or one of the keys, to providing the level of reassurance that is wanted.[38]

As yet, few archivists know of the work of the Provenance Incubator Group of the W3C – the organization that develops standards for the Web – or are aware that this group has proposed a definition of provenance. According to the W3C, the provenance of a resource is 'a record that describes entities and processes involved in producing and delivering or otherwise influencing that resource'.[39] We may think that there is some confusion here between provenance itself and the documentation that describes provenance; some of us may not like this definition at all; but we cannot avoid the conclusion that defining provenance is no longer the sole prerogative of the archival profession, if indeed it ever was.

In 2011, one group of scientists wrote about integrity, authenticity, trust and provenance in the following terms:

> Integrity and authenticity of observation data ... lie at the heart of scientific research. Authenticity requires that any data or results presented are exactly what they are claimed to be. Integrity requires that the processes and transformations to which data have been subjected have not introduced any undisclosed distortion or bias or loss .... To a large extent, it is necessary to trust the reporter in order to have confidence in the integrity and authenticity of what is reported. But we also look for evidence that supports the conclusions we reach through the exercise of such trust .... Provenance information is necessary for information quality and trust.[40]

Two years earlier, another group of computer science researchers said simply that we have become 'increasingly dependent on digital information' and that 'to be able to trust a piece of information, we need to know its provenance'.[41]

The message promoted by these research groups is that, while data can and should be reusable and reused, each piece of data should carry with it some evidence of its history and its original context, to help those who encounter it to form a judgement about its trustworthiness. The work that has already been done on this by computing specialists is quite extensive, and in general this work is a response to concerns very similar to those of the archivists mentioned earlier, although most of the accounts in the published literature of data management show little or no awareness of understandings of provenance in archival science. The existence of a number of competing research teams means that several different models, ontologies or vocabulary sets for representing provenance can now be found within the data management community; these include the Provenance Vocabulary Model, the Provenir Upper Ontology, the SWAN Provenance Vocabulary and, the best known and perhaps the most generic, the Open Provenance Model.[42] This last is intended – or so its designers claim – to support representation of the provenance of

non-digital as well as digital resources; it is also designed to allow provenance information to be exchanged between systems.[43]

As noted above, much work in information technology has been concerned with the provenance of *data*, and from an archival viewpoint this work sometimes displays a rather limited perspective. Provenance, one team of computing specialists has asserted, is 'information that helps determine the derivation history of a data product, starting from its original sources'.[44] Particularly in e-science, provenance is often associated with workflow, and is treated more or less as a synonym for 'derivation': it tracks the steps by which data have been derived or calculated from 'raw' scientific observations.

However, other computing experts take a wider view, and see provenance as embracing the origins, lineage and transmission history of any digital object.[45] In 2008, the Media Standards Trust and the Web Science Trust launched a 'Transparency Initiative' that looked at ways of making the provenance of news items more transparent online.[46] Some of the most recent work in information technology has sought to build provenance architectures to enhance trust in the domains of linked data and the Semantic Web[47] and also to extend such architectures to the world of cloud computing.[48] According to a computer scientist at Edinburgh University, 'provenance is the foundation for any reasonable model of privacy and trust ... on the Semantic Web'. A team at Harvard University has set out to make the case that provenance is crucial for information stored in the cloud and to argue that mechanisms for capturing and storing provenance metadata should be incorporated as a core feature in cloud services.[49]

Other computer specialists have shown an interest in provenance, or context, for reasons that initially may seem less closely linked to concerns about trust. For example, some specialists work on developing what are called 'context aware systems': systems that are designed to take account of users' actions, states or surroundings and to enable computing devices to adapt accordingly. In many cases, the aim of such systems is to minimize the need for explicit user interaction with a computer, and they are often associated with healthcare settings and with assisted living for the elderly or those with disabilities.[50] In 'pervasive' or 'ubiquitous' computing, sensor technology is applied to everyday objects to make them aware of, and responsive to, the contexts in which we use them.[51] Other research has set out to develop context aware systems that can be employed in the workplace. Many of these systems seek awareness of documentary contexts and workplace activities, as a means of providing support for users of digital resources, in what is often referred to as 'activity-based computing'. Some of them aim to allow users to interact with computers at the level of the activity, by employing interfaces that correspond to work processes, as an alternative to the application-centric or document-centric views that prevail in most mainstream user interfaces.[52] Others seek to capture and preserve information about users' activities, and thus implicitly or explicitly to maintain information about the provenance of the records that users create, by documenting at least some aspects of the immediate context or contexts in which creation of records occurred. Within the last decade, both in Europe and in the USA, researchers have proposed 'semantic' file systems based on 'context awareness' or 'provenance relationships'.[53] Another European research project, NEPOMUK, has sought to relate digital objects to agents, tasks and other contextual entities in order to develop a 'social semantic desktop'.[54]

Capturing context

As might be expected, not all these initiatives have developed beyond the prototype stage, and researchers in these projects have viewed context and provenance relationships from a

computing rather than an archival science perspective, typically emphasizing their role as data management or access mechanisms. Nevertheless, the provenance research undertaken by computer scientists is potentially of great interest to archivists, not least because of the scale of the challenge that archivists face in capturing and maintaining contextual information in the world of digital records. The labour-intensiveness of manual capture of metadata has long been known, and most archival institutions have large backlogs of unprocessed materials. If creators and users can contribute metadata, the burden of capture is shared, but attempts by records managers to persuade users of electronic records management systems to supply contextual metadata at the point of creation have often met considerable resistance. Even free-form tagging, which normally requires less effort than more formal methods of metadata capture, may not be adopted by a sufficient number of users. To achieve the volumes of metadata that may be needed, archivists and records managers will almost certainly have to supplement manual capture with artificial intelligence systems that automatically analyse the form, content and context of objects and populate descriptive systems accordingly.

The good news for archivists is that computer scientists who seek to document provenance face similar resource challenges and have begun to seek ways of using technology to address them. Methods of automatically harvesting certain kinds of contextual information are under investigation and could prove greatly valuable in archival work. Again, much computer science research in this area is concerned with automated capture of the provenance of *data*,[55] but there are also research projects that seek automated means of capturing knowledge about activities in the workplace[56] or the context of personal life.[57]

The systems developed in these projects typically monitor users' interactions with computers and data resources, gather information about these interactions and then process this information to make inferences about the context of users' activities and attempt to discover relationships among the tasks that users perform. The approach varies slightly from one project to another, but the American CAAD project (for example) employs a pattern mining algorithm to detect structures in a user's workflow that are believed to 'encode the content and context of the user's work activities'.[58] The Austrian UICO project employs sensors to observe user interaction behaviour and then uses machine learning techniques to populate a context model.[59]

While these systems aim to capture contextual events dynamically as they occur, there are also initiatives seeking to extract contextual information retrospectively in areas such as digital video curation and e-mail analysis.[60] These too use a range of computational processes drawn from the fields of data mining and artificial intelligence, including semantic indexing, social network analysis, similarity matching and visual analytics.

In general, archivists and records managers are little aware of this research, which in any case is still in its infancy at present. Inference tools in particular have proved very challenging to implement. The relevance of these systems to archival work, however, is clear, and in the longer term we can expect to see new tools that archivists and records managers can use, if technological frameworks are in place to allow and support their adoption. Where documentation of provenance and context is concerned, automated solutions are likely to expedite what are currently time-consuming manual processes.

Of course, some notes of caution are necessary. The contexts of records include not only the actions in which records are involved and the persons connected with them, but also the wider environments in which records are created and maintained. These wider environments include, but are in no way limited to, the functions of workgroups, organizations and nuclear and extended families; they also extend to the broad societal,

legal, cultural and physical contexts in which individuals, families, partnerships, workgroups, communities and organizations operate. Nor is context confined to a moment or moments in time; behind every one of its aspects lies a complex web of prior events, occurrences, actions and situations, an evolving sequence of historical contexts for what we perceive as the contexts of the present day. As South African archivist Verne Harris has said, 'there is no end to a search through this terrain';[61] context is infinite, and every context has contexts of its own. In practice, if archivists and archival institutions seek to document context, they have to decide what levels of context are most relevant to their needs or the needs of their users; this means making decisions about what to emphasize and what to ignore, decisions that will inevitably privilege some aspects of context above others.

Even though archivists may not feel it necessary or feasible to document what philosopher John Searle called 'the background' – the store of common abilities and knowledge about the world which, rightly or wrongly, most or all of us are tacitly assumed to share[62] – it is likely that they will still feel a need to capture some wider contextual information beyond an immediate low-level context of names, dates and places. While they may have confidence in the ability of computer applications to gather the information that a certain record was created by a certain person at a certain time, and perhaps also that this person was at that moment working on a certain project with certain named co-workers at a certain location, the further they want to extend their documentation of provenance beyond these immediate facts into the realm of broad societal, cultural and physical contexts, the less confident they may become about the capabilities of the computer to detect such broad contexts and also about their own ability to determine an appropriate mode of description. Archivists have traditionally believed that, if sufficient care is taken, they can create descriptions that accurately and impartially represent the things they set out to describe;[63] but there is now increasing recognition that definitive descriptions of ambient contexts are impossible, and that the interpretation of complex conceptual entities and the delineation of their boundaries will always be open to dispute. If humans cannot hope to compile descriptions that represent every nuance and subtlety of context, still less can we expect a computer to create definitive or richly nuanced descriptions on our behalf. Nor can we find a solution to these dilemmas by seeking refuge in the kind of computer science definition that identifies context as merely those 'elements of a user's environment which the computer knows about'.[64] Automated systems can only aspire to capture those aspects of context to which they are directly or indirectly exposed, and we must acknowledge that they are unlikely to supply all kinds of contextual information that future researchers may require.

Nevertheless, recognition that representations of context are partial and imperfect does not mean that such representations are useless. Whatever reservations we may have about the limitations of some of the current work in computer science, we can assume that future tools will almost certainly be more powerful. It seems reasonable to foresee a time when artificial intelligence will relieve archivists of many of the more mundane aspects of generating contextual documentation, thereby allowing archival institutions to concentrate resources on those aspects that cannot easily or effectively be automated. If the arguments put forward here are correct, such documentation forms a basis for enhancing the transparency of records, and hence for evaluating their trustworthiness.

Transparency, trust and provenance

What, then, are we to make of Weinberger's claim that transparency has replaced objectivity? Some computer science researchers have supported the suggestion that

centralized authority structures are no longer effectual. According to James Cheney and his colleagues,

> historically, databases … were trusted because they were under centralized control: it was assumed that trustworthy and knowledgeable people were responsible for the integrity of data …. Today, data is often made available on the Internet with no centralized control over its integrity …. Because information sources … vary widely in terms of quality, it is essential to provide provenance and other context information which can help end users judge whether query results are trustworthy.[65]

The connections between provenance, transparency and trust have also been widely asserted in the computing literature. A recent computer science research paper was entitled 'Trust and Provenance: You Can't Have One Without The Other'.[66] Luc Moreau, a computer scientist at the University of Southampton, has contended that 'provenance provides the necessary evidence which makes systems transparent'.[67] At Harvard, Uri Braun and his colleagues claimed that automatic capture systems provide 'a level of transparency … not found in more conventional provenance systems',[68] presumably because they do not depend on direct human input. At the Rensselaer Polytechnic Institute in New York State, e-science researchers Alvaro Graves and his colleagues affirmed that the provenance of data 'plays an important role in transparency', which in turn they identified as 'the only valid source of trust in science'.[69]

In 2012, in what appears to be the first study of the data provenance movement published in an archival journal, Kathleen Fear and Devan Ray Donaldson reported on their tests of computer scientists' claims that provenance plays a part in the processes users employ when attempting to determine the credibility of data sets.[70] Fear and Donaldson found that statements about provenance were important to the users they studied, but were not the only resource these users employed in reaching judgements about credibility. The users also relied on their assessment of the data content, on their own experience and pre-existing knowledge of data producers and on corroborative interrelationships between records. They found each of these factors useful, but none was sufficient in isolation; optimum judgements depended on combining them all.[71]

These findings, although scarcely surprising, are of interest in a number of ways. They reaffirm the role of provenance in decisions about trust, but they also remind us of the personal nature of those decisions. The computer science literature often asserts or assumes that provenance information can be encoded in a manner that permits software to calculate precise and seemingly objective 'trust ratings' or 'trust values',[72] but the users in Fear and Donaldson's study each formed their own judgements based on personal knowledge and individual evaluation. Trust is a matter of choice; I may choose to trust a particular resource and you may choose not to trust it, even when we are presented with the same statements about its provenance. An 'Oh yeah?' button that offers a formalized 'trust rating' for the resource seems likely to be ridiculed by many thoughtful users; even if it does not merely reflect someone else's opinion, at best the rating will be derived from a software program based on assumptions that the user may or may not share.[73] Universal agreement on what can be considered trustworthy is unlikely ever to be achieved.

Diplomatic science reminds us that trust in records operates at various levels: a record, or something that purports to be a record, can be judged in terms both of its authenticity (do I believe that it is really what it claims to be?) and of its accuracy or reliability (do I believe that it correctly represents the facts, actions or events that it claims to represent?).[74] Of course, all of these are contested notions. Records management literature often tells us that reliable records must be both full and accurate;[75] but in reality, once we move beyond simple statements about things such as dates, times and geographical

locations into the realm of discursive description, there is scope for a variety of ways of representing actions or events, and few now accept the possibility of single definitive accounts whose accuracy cannot be challenged and whose completeness is incontrovertible. In practice, I may choose to trust the account in the record I have to hand while recognizing that other accounts are possible and that I could perhaps have confidence in them also. Sometimes I may decide that I cannot wholeheartedly place my trust in the contents of any of the available records. To add a further tier of complexity, if I have a record of assertions made by interested parties, I may choose to trust that the record tells me what those parties said, even if I feel that the propositions they made are open to doubt.[76]

Authenticity seems to be a different matter; even if I lack confidence in the contents of the record, I may still believe it to be authentic in the sense that it has not been forged, corrupted or tampered with. Because authenticity can be endangered by weaknesses in transmission and preservation procedures, if I wish to make a judgement about the authenticity of a record I am likely to want access to contextual information about the record's custodial history and its adventures over time. In archival discourse, the term 'provenance' has traditionally been confined to understandings of a record's origins and the contexts of its creation, but many archivists now argue that it should also embrace information about custodial history, thus aligning archival notions of provenance more closely to those of the library and museum communities.

Inevitably, however, a need for information about provenance or context will incur a further trust-related decision. If I want to use any kind of information about provenance to help me form judgements about the trustworthiness of records, I will have to decide whether I feel able to trust the provenance information itself. Whether it is information about the origins of records or information about their custodial history, what is sought is knowledge of the past; but none of us can experience the past directly in the present; we can only view it through representations, and we must consider how far we are willing to trust those representations. Perhaps I may have some personal memory of the circumstances in which a record was created – a mental representation of my own that I can draw on – but more commonly I will be dependent on representations of provenance that have been created by others. What I will need to recognize is that the representations we call provenance metadata are also records in their own right: they are records of statements or assertions that people have made about past situations and events, or records of those situations and events that have been captured by a computer. They may be no more, or no less, trustworthy than any other records. The need for confidence about authenticity or reliability applies as much to metadata as to the resources they refer to or describe.

As with any judgement about trust, an evaluation of provenance metadata is likely to rest on a number of building blocks: do the metadata look right? do they feel right? do they match up to our prior knowledge of the sorts of things that experience tells us we can trust? are they corroborated by other sources of information to which we have access? If we are technically inclined, or concerned that guilty parties who falsify a record may also want to tamper with its provenance metadata in order to cover their tracks, we may perhaps look to cryptographic or forensic techniques to provide reassurance that statements about provenance have not been corrupted. But it also seems necessary to consider 'the provenance of the provenance statement itself': who provided the details of provenance, when they did so and under what circumstances. A recent European project on computer science workflow referred to this as *meta-provenance*, which it defined as a 'meta-statement about a provenance statement or set of provenance statements'.[77] As the project

report noted, 'without any knowledge of the provenance of provenance information, we risk drawing conclusions about trustworthiness based on … untrustworthy data'.[78]

This issue may be especially important if the description of records and archives in cyberspace is opened up to user contributions, as many archivists have advocated in recent years.[79] Arguably, when 'provenance statements can be made by anyone, at any time',[80] there may be a particular need to consider how far we feel able to trust their creators. If the details provided to us are the 'wisdom of the crowd', we may want to know how many people have contributed to this wisdom, and what kind of people they were. If only one or two people have contributed, it may be even more critical to discern their identities, if we want to assess our confidence in the contributions they made. Are they people we know? Are they people whose reputation is known to us? Do we have other evidence that prompts us to trust them? As Donovan Artz and Yolanda Gil have observed, in web or web-like environments where 'anyone can say anything about anything … we need to be able to understand where we are placing our trust.'[81]

Conclusion

According to the W3C, 'provenance provides a critical foundation for assessing authenticity [and] enabling trust'.[82] Provenance is undoubtedly an important ingredient in evaluating trust in records and archives, in digital as much as in analogue realms, but we must conclude that merely being supplied with information about provenance is not enough. Insofar as there are individuals or institutions who provide us with that information, we have to decide how far we trust what those people or institutions tell us, and perhaps also how far we trust the particular environment in which we learn what they say. There may seem to be a recursive argument here, an apparently endless need for metadata about metadata, but a 'law of diminishing returns' is likely to operate and ultimately we will have to rely on our personal judgement. Regardless of whether the statements we read about provenance or meta-provenance are provided by a single source such as an archivist or by a 'crowd' of sources, we must use our knowledge or our perception of the creators of those statements and the publishers of the databases or websites where the statements are made, and our level of confidence that creators and publishers are who they say they are, in deciding whether we are willing to trust them. If the statements we read have been captured automatically by a computer application, we must similarly rely on our judgement of the application and those who wrote it. We must also accept that we are fallible and sometimes will make the wrong choice: we will opt to trust when it would be better not to do so, or perhaps we will mistrust someone or something that we could have found trustworthy.

Finally, as we examine this territory, we find that issues of professional authority and expertise resurface. Despite the current fashion for disparaging professionalism, I would like to suggest that the professional archivist and the archival institution still have key roles to play in securing trust. Reputation is important here; just as we do may choose to do business with Amazon rather than with unknownbookseller.com, because we trust the reputation of an established name, so when we come to assess statements about the provenance or context of records and archives we may be more likely to trust those made by people or institutions we perceive as reputable and competent. In a move that is sometimes labelled as 'trust transfer',[83] we may find that we are willing to trust a largely unknown entity on the basis of its association with a known source in which we have fuller confidence. Even in a world where trust depends largely on transparency, there is still a place for the expert voice, although we must recognize that experts cannot be seen as

infallible providers of objective information. Like other experts, archivists and record-keepers now have to earn the trust that is reposed in them; they can no longer merely assume it.

Notes

1. The workshop was held at the Peter Wall Institute at the University of British Columbia, Vancouver, Canada, in September 2012. See http://www.digitaltrust.pwias.ubc.ca/.
2. Both the address and the published paper owe much to conversations with Fiona Cormack, Vicki Lemieux, Elizabeth Shaffer and David Thomas; I am very grateful for their comments and suggestions.
3. Trust is a famously elusive and contentious concept, but for the purposes of this paper it can be broadly characterized as a state of confidence in some person or thing, or an expectation that some person or thing will prove worthy of confidence. The *Oxford English Dictionary* offers definitions of trust as 'confident expectation of something' and 'confidence in or reliance on some quality or attribute of a person or thing, or the truth of a statement'. For fuller discussion of different understandings of trust, see Blomqvist, "Many Faces of Trust" and McKnight and Chervany, "Meanings of Trust."
4. Locke, *Essays on the Law of Nature*, 213.
5. Cf. Cvetkovich and Löfstedt, *Social Trust and the Management of Risk*. See also Riegelsberger, Sasse and McCarthy, "Mechanics of Trust," 385.
6. There is a substantial literature on the topic of 'institutional trust' and how it may or may not differ from other forms of trust. See, for example, Lahno, "Institutional Trust."
7. Löfstedt and Horlick-Jones, "Environmental Regulation in the UK," 74–6, 81 and Viklund, "Risk Policy," 14, 23–4. See also European Commission Directorate-General for Communication, *Eurobarometer 76*, 19.
8. See, for example, Martin Vander Weyer, "Never Mind the Banks" (*The Spectator*, July 14, 2012).
9. Fukuyama, *Trust: The Social Virtues*, especially 7–17.
10. Hosking, *Trust: Money, Markets*, 60–2.
11. Paul Vallely, "How to Change the Bollinger Mindset" (*The Independent on Sunday*, July 1, 2012) and Michael Buerk, "Moral Maze" (*BBC Radio 4*, July 4, 2012). Cf. Power, *Audit Society*, 120–1.
12. Cofta, *Trust, Complexity and Control*, 122–3.
13. Lynch, "Authenticity and Integrity," 33.
14. Nhan, Kinkade and Burns, "Finding a Pot of Gold"; Anon., "An Old Swindle Revived" (*The New York Times*, March 20, 1898) and Morrish, "Fraud in Various Forms," 594.
15. Perhaps the most obvious example is the online encyclopaedia Wikipedia, which offers a greater quantity of information than any paper encyclopaedia has ever provided, but is open to incompetent editing and malicious vandalism (and, of course, to subsequent correction). Like its smaller and more overtly slanted competitors such as Conservapedia, Wikipedia has also been suspected of political bias. See Denning et al., "Wikipedia Risks," 152; Dalby, *World and Wikipedia*, 72–81.
16. Lynch, "Authenticity and Integrity," 33. Cf. Kelton, Fleischmann and Wallace, "Trust in Digital Information," 363.
17. Berners-Lee, "Cleaning Up the User Interface."
18. Statistics Canada, "Statistics Canada's Quality Assurance."
19. Weinberger, "Transparency Is the New Objectivity."
20. For an account (from a broadly legal perspective) of issues relating to trust and record-keeping in the 'cloud', see Duranti and Rogers, "Trust in Digital Records." Wider issues of trust in records and archives formed the theme of a FARMER (Forum for Archives and Records Management Education and Research) conference at Oxford in July 2010; selected papers from this conference were published in a special issue of *Archival Science* (11, no. 3–4) in 2011.
21. See MacNeil, *Trusting Records*.
22. For an examination of the limited acceptance of digital signature technology, see Blanchette, *Burdens of Proof*. According to Blanchette, cryptographers 'have tended to assume that the properties of cryptographic objects will translate transparently into complex social and institutional settings' (p. 127), but have glossed over 'the enormous challenges inherent in

turning cryptographic constructs into … technologies able to perform in the context of users' everyday lives' (p. 84), with the result that 'the business case for … encryption and digital signatures failed to carry the marketplace' (p. 5).

23. MacNeil, "Trusting Description," 90–1 and MacNeil, "Trust and Professional Identity," 180–1.
24. Gilliland, "Neutrality, Social Justice," 196–7. Cf. Borland, "Trust and the Records Professional," 2, 6 and Millar, *Archives*, 46.
25. Weinberger, "Transparency is the New Objectivity."
26. Dollar, *Archival Theory and Information Technologies*, 48.
27. Cook, *Management of Information*, 102; McKemmish, "Introducing Archives and Archival Programs," 12 and MacNeil, "Picking Our Text," 271–2. Cf. Jenkinson, *Manual of Archive Administration*, 83.
28. See, for example, Bastian, *Owning Memory*, 81–3; Millar, "Death of the Fonds," 12–4; Nesmith, "Seeing Archives," 35–6.
29. See, for example, Horsman, "Wrapping Records in Narratives," 2 and Schwartz, "Archival Garden," 73.
30. Roper, "Development of the Principles" and Procter, "Life Before Jenkinson," 149–51.
31. Muller, Feith and Fruin, *Handleiding voor het Ordenen*.
32. Blouin and Rosenberg, *Processing the Past*, 203.
33. Lee, "Collecting the Externalized Me," 208–9.
34. Mayer-Schönberger, *Delete: The Virtue of Forgetting*, 78, 90.
35. Berg and Goorman, "Contextual Nature of Medical Information"; the quotations are from pp. 52 and 54.
36. Archivists have traditionally attempted to preserve knowledge of logical interrelationships by means of physical aggregations and fixed arrangements, but in Yeo, "Bringing Things Together," I argued that 'no single arrangement captures all the interconnections that might be of interest' (p. 65) and that other approaches (using descriptive metadata and granular relational models or ontologies) can now be recognized as more powerful and effective.
37. Mayer-Schönberger, *Delete: The Virtue of Forgetting*, 79.
38. See, for example, Dai et al., "Approach to Evaluate Data"; Cheney, Chiticariu and Tan, "Provenance in Databases" and Moreau, "Foundations for Provenance." For a useful introduction to the world of big data and a discussion of its implications for the wider society, see Mayer-Schönberger and Cukier, *Big Data*.
39. W3C Provenance Incubator Group, "Provenance XG Final Report."
40. Wf4ever Advanced Workflow Preservation Technologies for Enhanced Science, "Workflow Integrity and Authenticity," 8, 10.
41. Hasan, Sion and Winslett, "Secure Provenance," 12.
42. Wf4ever Advanced Workflow Preservation Technologies for Enhanced Science, "Workflow Integrity and Authenticity," 12 and Ding et al., "Reflections on Provenance Ontology," 199–200.
43. Moreau et al., "Open Provenance Model," 2.
44. Simmhan, Plale and Gannon, "Survey of Data Provenance," 31.
45. Hasan, Sion and Winslett, "Secure Provenance," 13.
46. See http://mediastandardstrust.org/projects/transparency-initiative/.
47. Omitola, Gibbins and Shadbolt, "Provenance in Linked Data Integration" and Halpin, "Provenance."
48. Muniswamy-Reddy et al., "Provenance for the Cloud" and Imran and Hlavacs, "Provenance in the Cloud."
49. Halpin, "Provenance," unpaginated and Muniswamy-Reddy et al., "Provenance for the Cloud," unpaginated.
50. Davies, Siewiorek and Sukthankar, "Activity-Based Computing."
51. Greenfield, *Everyware* and Schmidt et al., "Interacting with 21st-Century Computers."
52. Bardram, Bunde-Pedersen and Soegaard, "Support for Activity-Based Computing" and Voida, Mynatt and Edwards, "Re-framing the Desktop Interface." These computer science approaches have many similarities to those advocated within the archival community in the 1990s by David Bearman and John McDonald (Bearman, "Item Level Control," 221–7 and McDonald, "Towards Automated Record Keeping").
53. Karypidis and Lalis, "Automated Context Aggregation" and Leung, Parker-Wood and Miller, "Copernicus."

54. Grimnes, Sauermann and Bernardi, "Personal Knowledge Workbench" and Riss et al., "Knowledge Work Support." An earlier project along similar lines, but with perhaps more emphasis on the workplace and on the flexible and adaptable nature of the technology, is described in Moran, Cozzi and Farrell, "Unified Activity Management."
55. Braun et al., "Issues in Automatic Provenance"; Barga and Digiampietri, "Automatic Generation of Workflow" and Wombacher and Huq, "Towards Automatic Capturing."
56. Dragunov et al., "TaskTracer" and Brdiczka, "From Documents to Tasks."
57. Karypidis and Lalis, "Automated Context Aggregation."
58. Rattenbury and Canny, "CAAD," 687.
59. Rath, Devaurs and Lindstaedt, "UICO." Other projects that have sought to develop tools for automated or semi-automated detection and capture of contextual activity include UMEA (Kaptelinin, "UMEA"), TaskTracer (Dragunov et al., "TaskTracer") and Smart Desktop (Lowd and Kushmerick, "Using Salience to Segment").
60. Shah, "Mining Contextual Information"; Perer and Shneiderman, "Beyond Threads"; Dredze, Lau and Kushmerick, "Automatically Classifying Emails"; Esteva et al., "Finding Narratives of Activities"; Mayer, Neumayer and Rauber, "Interacting with (Semi-) Automatically Extracted Context" and Kang et al., "Making Sense of Archived E-Mail."
61. Harris, *Exploring Archives*, 83.
62. Searle, *Consciousness and Language*, 153–4, 196–7. Several studies have noted that users of archives sometimes feel little or no requirement for knowledge of previous contexts (see, for example, Craven, "From the Archivist's Cardigan," 19–20); but it would probably be more correct to say that such users seek no further explicit presentation of context because their prior 'background' capacities and knowledge already suffice for their needs. Genealogical users, for example, are often thought to want specific information and to lack interest in the wider contexts of the records they consult, but if information in baptism registers, wills or apprenticeship indentures appears to need no elucidation it is only because most users already possess some contextual understanding of the societal practices these records represent. Context knows no limits, but as noted by Artz and Gil, "Survey of Trust," 58, 'humans … bring to bear vast amounts of knowledge about the world they live in', and arguably this allows us to assume that certain aspects of context can be left undocumented.
63. See, for example, the account of description in Haworth, "Archival Description."
64. Wootten and Rana, "Recording the Context," 45.
65. Cheney, Chiticariu and Tan, "Provenance in Databases," 380.
66. Janowicz, "Trust and Provenance."
67. Moreau, "Foundations for Provenance," 68.
68. Braun et al., "Issues in Automatic Provenance," 171.
69. Graves, Lebo and McCusker, "Provenance and Trust in E-Science."
70. Fear and Donaldson, "Provenance and Credibility." These authors reported similar conclusions in Donaldson and Fear, "Provenance, End-User Trust and Reuse."
71. Fear and Donaldson, "Provenance and Credibility," 327–33 and Donaldson and Fear, "Provenance, End-User Trust and Reuse," unpaginated.
72. See, for example, Golbeck and Mannes, "Using Trust and Provenance"; Dai et al., 'Approach to Evaluate Data" and Chapman, Blaustein and Elsaesser, "Provenance-based Belief."
73. The computer industry has attempted to respond to concerns about trustworthiness in e-commerce by providing logos or 'trust seals' that purport to offer assurance that a website is reputable. These devices are usually intended to convey information that an online vendor has subscribed to a code of conduct administered by a certification authority. An anonymous reviewer for *Archives & Records* has drawn my attention to a recent paper in a computer science journal (Kirlappos, Sasse and Harvey, "Why Trust Seals Don't Work"), which indicated that these devices are easily forged and that users frequently overlook them or find them puzzling. Kirlappos, Sasse and Harvey suggested that the answer probably lies in 'automatic verification' of trustworthiness (p. 308), but 'verification' remains a highly loaded term. MacNeil, "Trust and Professional Identity," has given us a timely reminder that 'the trustworthiness of records is socially negotiated, [and] historically situated' (p. 187). Archivists and the wider public may not always share computer scientists' confidence in the possibilities of objective authentication and of constructing infallible solutions that wholly depend on software programming.
74. Duranti, "Reliability and Authenticity."

75. ISO 15489–1, sec. 7.2.3; State Records Authority of New South Wales, *Standard on Full and Accurate Records*; Queensland State Archives, *Creating Full and Accurate Records*.
76. Yeo, "Representing the Act," 105–6.
77. Wf4ever Advanced Workflow Preservation Technologies for Enhanced Science, "Workflow Integrity and Authenticity," 15.
78. Ibid., 16.
79. Anderson and Allen, "Envisioning the Archival Commons"; Evans, "Archives of the People" and Huvila, "Participatory Archive."
80. Wf4ever Advanced Workflow Preservation Technologies for Enhanced Science, "Workflow Integrity and Authenticity," 15.
81. Artz and Gil, "Survey of Trust," 58.
82. See the W3C website at http://www.w3.org/2005/Incubator/prov/wiki/What_Is_Provenance.
83. McEvily, Perrone and Zaheer, "Introduction to the Special Issue," 2.

References

Anderson, Scott R., and Robert B. Allen. "Envisioning the Archival Commons." *American Archivist* 72, no. 2 (2009): 383–400.

Artz, Donovan, and Yolanda Gil. "A Survey of Trust in Computer Science and the Semantic Web." *Journal of Web Semantics* 5, no. 2 (2007): 58–71.

Bardram, Jakob E., Jonathan Bunde-Pedersen, and Mads Soegaard. "Support for Activity-Based Computing in a Personal Computing Operating System." *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press, 2006.

Barga, Roger S., and Luciano A. Digiampietri. "Automatic Generation of Workflow Provenance." *Lecture Notes on Computer Science* 4145 (2006): 1–9.

Bastian, Jeannette A. *Owning Memory: How a Caribbean Community Lost Its Archives and Found Its History*. Westport, CT: Libraries Unlimited, 2003.

Bearman, David. "Item Level Control and Electronic Recordkeeping." *Archives and Museum Informatics* 10, no. 3 (1996): 195–245.

Berg, Marc, and Els Goorman. "The Contextual Nature of Medical Information." *International Journal of Medical Informatics* 56 (1999): 51–60.

Berners-Lee, Tim. "Cleaning Up the User Interface." 1997. http://www.w3.org/DesignIssues/UI.html

Blanchette, Jean-François. *Burdens of Proof: Cryptographic Culture and Evidence Law in the Age of Electronic Documents*. Cambridge, MA: MIT Press, 2012.

Blomqvist, Kirsimarja. "The Many Faces of Trust." *Scandinavian Journal of Management* 13, no. 3 (1997): 271–286.

Blouin, Francis X., and William G. Rosenberg. *Processing the Past: Contesting Authority in History and the Archives*. New York: Oxford University Press, 2011.

Borland, Jennifer. "Trust and the Records Professional." 2009. http://www.armaedfoundation.org/pdfs/JBorland_ScholarshipEssay.pdf

Braun, Uri, Simson Garfinkel, David A. Holland, Kiran-Kumar Muniswamy-Reddy, and Margo I. Seltzer. "Issues in Automatic Provenance Collection." *Lecture Notes on Computer Science* 4145 (2006): 171–183.

Brdiczka, Oliver. "From Documents to Tasks: Deriving User Tasks from Document Usage Patterns." *IUI 2010: Proceedings of the 15th International Conference on Intelligent User Interfaces*. New York: ACM Press, 2010.

Chapman, Adriane, Barbara Blaustein, and Chris Elsaesser. "Provenance-Based Belief." 2010. http://www.mitre.org/work/tech_papers/2010/09_5315/09_5315.pdf

Cheney, James, Laura Chiticariu, and Wang-Chiew Tan. "Provenance in Databases: Why, How, and Where." *Foundations and Trends in Databases* 1, no. 4 (2009): 379–474.

Cofta, Piotr. *Trust, Complexity and Control*. Chichester: John Wiley & Sons, 2007.

Cook, Michael. *The Management of Information from Archives*, 2nd ed. Aldershot: Gower, 1999.

Craven, Louise. "From the Archivist's Cardigan to the Very Dead Sheep: What are Archives? What are Archivists? What do They Do?" In *What are Archives? Cultural and Theoretical Perspectives: A Reader*, edited by L. Craven. Aldershot: Ashgate, 2008.

Cvetkovich, George, and Ragnar E. Löfstedt, eds. *Social Trust and the Management of Risk*. London: Earthscan Publications, 1999.

Dai, Chenyun, Dan Lin, Elisa Bertino, and Murat Kantarcioglu. "An Approach to Evaluate Data Trustworthiness Based on Data Provenance." *Lecture Notes on Computer Science* 5159 (2008): 82–98.

Dalby, Andrew. *The World and Wikipedia*. Draycott: Siduri Books, 2009.

Davies, Nigel, Daniel P. Siewiorek, and Rahul Sukthankar. "Activity-Based Computing." *Pervasive Computing* 7, no. 2 (2008): 20–21.

Denning, Peter, Jim Horning, David Parnas, and Lauren Weinstein. "Wikipedia Risks." *Communications of the ACM* 48, no. 12 (2005): 152.

Ding, Li, Jie Bao, James R Michaelis, Jun Zhao, and Deborah L. McGuinness. "Reflections on Provenance Ontology Encodings." *Lecture Notes on Computer Science* 6378 (2010): 198–205.

Dollar, Charles M. *Archival Theory and Information Technologies: The Impact of Information Technologies on Archival Principles and Methods*. Macerata, Italy: University of Macerata, 1992.

Donaldson, Devan Ray, and Kathleen Fear. "Provenance, End-User Trust and Reuse: An Empirical Investigation." 2011. http://static.usenix.org/events/tapp11/tech/final_files/Donaldson.pdf

Dragunov, Anton N., Thomas G. Dietterich, Kevin Johnsrude, Matthew McLaughlin, Lida Li, and Jonathan L. Herlocker. "TaskTracer: A Desktop Environment to Support Multi-tasking Knowledge Workers." *IUI '05: Proceedings of the 10th International Conference on Intelligent User Interfaces*. New York: ACM Press, 2005.

Dredze, Mark, Tessa Lau, and Nicholas Kushmerick. "Automatically Classifying Emails into Activities." *IUI '06: Proceedings of the 11th International Conference on Intelligent User Interfaces*. New York: ACM Press, 2006.

Duranti, Luciana. "Reliability and Authenticity: The Concepts and their Implications." *Archivaria* 39 (1995): 5–10.

Duranti, Luciana, and Corinne Rogers. "Trust in Digital Records: An Increasingly Cloudy Legal Area." *Computer Law & Security Review* 28 (2012): 522–531.

Esteva, Maria, Weijia Xu, Jaya Sreevelsan-Nair, Ashwini Athalye, and Merwan Hadethe. "Finding Narratives of Activities through Archival Bond in Electronically Stored Information." 2009. http://web.archive.org/web/20100722172246/http://www.law.pitt.edu/DESI3_Workshop/Papers/DESI_III.Esteva-Xu-Nair.pdf

European Commission Directorate-General for Communication. "Eurobarometer 76: Public Opinion in the European Union: First Results." 2011. http://ec.europa.eu/public_opinion/archives/eb/eb76/eb76_first_en.pdf

Evans, Max J. "Archives of the People, by the People, for the People." *American Archivist* 70, no. 2 (2007): 387–400.

Fear, Kathleen, and Devan Ray Donaldson. "Provenance and Credibility in Scientific Data Repositories." *Archival Science* 12, no. 3 (2012): 319–339.

Fukuyama, Francis. *Trust: The Social Virtues and the Creation of Prosperity*. London: Penguin, 1996.

Gilliland, Anne. "Neutrality, Social Justice and the Obligations of Archival Education and Educators in the Twenty-first Century." *Archival Science* 11, no. 3-4 (2011): 193–209.

Golbeck, Jennifer, and Aaron Mannes. "Using Trust and Provenance for Content Filtering on the Semantic Web." 2006. http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-190/paper02.pdf

Graves, Alvaro, Tim Lebo, and Jim McCusker. "Provenance and Trust in E-Science." c.2010 . http://tw.rpi.edu/proj/portal.wiki/images/0/06/McCusker_Lebo_Graves_ProvenanceTrustSlides.pdf

Greenfield, Adam. *Everyware: The Dawning Age of Ubiquitous Computing*. Berkeley, CA: Peachpit Press, 2006.

Grimnes, Gunnar A., Leo Sauermann, and Ansgar Bernardi. "The Personal Knowledge Workbench of the NEPOMUK Social Semantic Desktop." *Lecture Notes in Computer Science* 5554 (2009): 836–840.

Halpin, Harry. "Provenance: The Missing Component of the Semantic Web for Privacy and Trust." 2009. http://ceur-ws.org/Vol-447/paper9.pdf

Harris, Verne. *Exploring Archives: An Introduction to Archival Ideas and Practice in South Africa*. Pretoria: National Archives of South Africa, 2000.

Hasan, Ragib, Radu Sion, and Marianne Winslett. "Secure Provenance: Protecting the Genealogy of Bits." 2009. http://static.usenix.org/publications/login/2009-06/openpdfs/hasan.pdf

Haworth, Kent M. "Archival Description: Content and Context in Search of Structure." In *Encoded Archival Description on the Internet*, edited by D. Pitti and W. Duff, New York: Haworth Press, 2002.

Horsman, Peter. "Wrapping Records in Narratives: Representing Context Through Archival Description." 2011. http://www.its-arolsen.org/fileadmin/user_upload/Dateien/Archivtagung/Horsman_text.pdf

Hosking, Geoffrey. *Trust: Money, Markets and Society*. Calcutta: Seagull Books, 2010.

Huvila, Isto. "Participatory Archive: Towards Decentralised Curation, Radical User Orientation, and Broader Contextualisation of Records Management." *Archival Science* 8, no. 1 (2008): 15–36.

Imran, Muhammad, and Helmut Hlavacs. "Provenance in the Cloud: Why and How?" 2012. http://www.thinkmind.org/download.php?articleid=cloud_computing_2012_5_20_20114.pdf

ISO 15489-1. *Information and Documentation: Records Management. Part 1: General*. International Organization for Standardization, 2001.

Janowicz, Krzysztof. "Trust and Provenance: You Can't Have One Without The Other." 2009. http://geog.ucsb.edu/~jano/trust_provenance.pdf

Jenkinson, Hilary. *A Manual of Archive Administration*, 2nd ed. London: Lund Humphries, 1937.

Kang, Hyunmo, Catherine Plaisant, Tamer Elsayed, and Douglas W. Oard. "Making Sense of Archived E-Mail: Exploring the Enron Collection with NetLens." *Journal of the American Society for Information Science and Technology* 61, no. 4 (2010): 723–744.

Kaptelinin, Victor. "UMEA: Translating Interaction Histories into Project Contexts." *CHI '03: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press, 2003.

Karypidis, Alexandros, and Spyros Lalis. "Automated Context Aggregation and File Annotation for PAN-based Computing." *Personal and Ubiquitous Computing* 11, no. 1 (2007): 33–44.

Kelton, Kari, Kenneth R. Fleischmann, and William A. Wallace. "Trust in Digital Information." *Journal of the American Society for Information Science and Technology* 59, no. 3 (2008): 363–374.

Kirlappos, Iacovos, M. Angela Sasse, and Nigel Harvey. "Why Trust Seals Don't Work: A Study of User Perceptions and Behavior." *Lecture Notes in Computer Science* 7344 (2012): 308–324.

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

Lee, Christopher A. "Collecting the Externalized Me: Appraisal of Materials in the Social Web." In *I, Digital: Personal Collections in the Digital Era*, edited by C. A. Lee. Chicago, IL: Society of American Archivists, 2011.

Leung, Andrew W., Aleatha Parker-Wood, and Ethan L. Miller. "Copernicus: A Scalable, High-Performance Semantic File System." 2009. http://ssrc.cse.ucsc.edu/Papers/ssrctr-09-06.pdf

Locke, John. *Essays on the Law of Nature*, edited by W. von Leyden. Oxford: Clarendon Press, 1954.

Löfstedt, Ragnar E., and Tom Horlick-Jones. "Environmental Regulation in the UK: Politics, Institutional Change and Public Trust." In *Social Trust and the Management of Risk*, edited by G. Cvetkovich and R. E. Löfstedt, London: Earthscan Publications, 1999.

Lowd, Daniel, and Nicholas Kushmerick. "Using Salience to Segment Desktop Activity into Projects." *IUI 2009: Proceedings of the 14th International Conference on Intelligent User Interfaces*. New York: ACM Press, 2009.

Lynch, Clifford. "Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust." *Authenticity in a Digital Environment*. Washington: Council on Library and Information Resources, 2000.

MacNeil, Heather. *Trusting Records: Legal, Historical, and Diplomatic Perspectives*. Dordrecht: Kluwer, 2000.

MacNeil, Heather. "Picking Our Text: Archival Description, Authenticity, and the Archivist as Editor." *American Archivist* 68, no. 2 (2005): 264–278.

MacNeil, Heather. "Trusting Description: Authenticity, Accountability, and Archival Description Standards." *Journal of Archival Organization* 7, no. 3 (2009): 89–107.

MacNeil, Heather. "Trust and Professional Identity: Narratives, Counter-Narratives and Lingering Ambiguities." *Archival Science* 11, no. 3–4 (2011): 175–192.

Mayer, Rudolf, Robert Neumayer, and Andreas Rauber. "Interacting with (Semi-) Automatically Extracted Context of Digital Objects." 2009. http://www.idi.ntnu.no/~neumayer/pubs/MAY09_ciao.pdf

Mayer-Schönberger, Viktor. *Delete: The Virtue of Forgetting in the Digital Age*. Princeton, NJ: Princeton University Press, 2009.
Mayer-Schönberger, Viktor, and Kenneth Cukier. *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray, 2013.
McDonald, John. "Towards Automated Record Keeping: Interfaces for the Capture of Records of Business Processes." *Archives and Museum Informatics* 11, no. 3–4 (1997): 277–285.
McEvily, Bill, Vincenzo Perrone, and Akbar Zaheer. "Introduction to the Special Issue on Trust in an Organizational Context." *Organization Science* 14, no. 1 (2003): 1–4.
McKemmish, Sue. "Introducing Archives and Archival Programs." In *Keeping Archives*, edited by J. Ellis. 2nd ed. Melbourne: D.W. Thorpe, 1993.
McKnight, D. Harrison, and Norman L. Chervany. "The Meanings of Trust." 1996. http://misrc.umn.edu/wpaper/WorkingPapers/9604.pdf
Millar, Laura A. "The Death of the Fonds and the Resurrection of Provenance: Archival Context in Space and Time." *Archivaria* 53 (2002): 1–15.
Millar, Laura A. *Archives: Principles and Practices*. London: Facet Publishing, 2010.
Moran, Thomas P., Alex Cozzi, and Stephen P. Farrell. "Unified Activity Management: Supporting People in E-Business." *Communications of the ACM* 48, no. 12 (2005): 67–70.
Moreau, Luc. "The Foundations for Provenance on the Web." 2010. http://eprints.soton.ac.uk/271691/
Moreau, Luc, Ben Clifford, Juliana Freire, Joe Futrelle, Yolanda Gil, Paul Groth, Natalia Kwasnikowska et al. "The Open Provenance Model Core Specification." 2010. http://eprints.soton.ac.uk/271449/
Morrish, R. "Fraud in Various Forms." *Police Journal* 3, no. 4 (1930): 589–600.
Muller, S., J. A. Feith, and R. Fruin. *Handleiding voor het Ordenen en Beschrijven van Archieven*. Groningen: Van der Kamp, 1898.
Muniswamy-Reddy, Kiran-Kumar, Peter Macko, and Margo Seltzer. "Provenance for the Cloud." 2010. http://static.usenix.org/event/fast10/tech/full_papers/muniswamy-reddy.pdf
Nesmith, Tom. "Seeing Archives: Postmodernism and the Changing Intellectual Place of Archives." *American Archivist* 65, no. 1 (2002): 24–41.
Nhan, Johnny, Patrick Kinkade, and Ronald Burns. "Finding a Pot of Gold at the End of an Internet Rainbow: Further Examination of Fraudulent Email Solicitation." *International Journal of Cyber Criminology* 3, no. 1 (2009): 452–475.
Omitola, Tope, Nicholas Gibbins, and Nigel Shadbolt. "Provenance in Linked Data Integration." 2010. http://linkeddata.future-internet.eu/images/e/eb/FIA2010_Provenance_in_the_Future_Internet.pdf
Perer, Adam, and Ben Shneiderman. "Beyond Threads: Identifying Discussion in Email Archives." 2005. http://hcil.cs.umd.edu/trs/2005-26/2005-26.pdf
Power, Michael. *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press, 1997.
Procter, Margaret. "Life Before Jenkinson: The Development of British Archival Theory and Thought at the Turn of the Twentieth Century." *Archives* 119 (2008): 140–161.
Queensland State Archives. "Creating Full and Accurate Records." 2008. http://www.archives.qld.gov.au/Recordkeeping/GRKDownloads/Documents/full_accurate.pdf
Rath, Andreas S., Didier Devaurs, and Stefanie N. Lindstaedt. "UICO: An Ontology-based User Interaction Context Model for Automatic Task Detection on the Computer Desktop." *Proceedings of the 1st Workshop on Context, Information and Ontologies*. New York: ACM Press, 2009.
Rattenbury, Tye, and John Canny. "CAAD: An Automatic Task Support System." *CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press, 2007.
Riegelsberger, Jens, M. Angela Sasse, and John D. McCarthy. "The Mechanics of Trust: A Framework for Research and Design." *International Journal of Human–Computer Studies* 62, no. 3 (2005): 381–422.
Riss, Uwe V., Olaf Grebner, Philip S. Taylor, and Ying Du. "Knowledge Work Support by Semantic Task Management." *Computers in Industry* 61 (2010): 798–805.
Roper, Michael. "The Development of the Principles of Provenance and Respect for Original Order in the Public Record Office." In *The Archival Imagination: Essays in Honour of Hugh A. Taylor*, edited by B. L. Craig. Ottawa: Association of Canadian Archivists, 1992.

Schmidt, Albrecht, Bastian Pfleging, Florian Alt, Alireza Sahami Shirazi, and Geraldine Fitzpatrick. "Interacting with 21st-Century Computers." *Pervasive Computing* 11, no. 1 (2012): 22–30.

Schwartz, Joan M. "The Archival Garden: Photographic Plantings, Interpretive Choices, and Alternative Narratives." In *Controlling the Past: Documenting Society and Institutions*, edited by T. Cook. Chicago, IL: Society of American Archivists, 2011.

Searle, John R. *Consciousness and Language*. Cambridge: Cambridge University Press, 2002.

Shah, Chirag. "Mining Contextual Information for Ephemeral Digital Video Preservation." *International Journal of Digital Curation* 4, no. 1 (2009): 175–192.

Simmhan, Yogesh L., Beth Plale, and Dennis Gannon. "A Survey of Data Provenance in E-Science." *SIGMOD Record* 34, no. 3 (2005): 31–36.

State Records Authority of New South Wales. "Standard on Full and Accurate Records." 2004. http://www.records.nsw.gov.au/documents/recordkeeping-standards/Standard%20No%20%207%20-%20Full%20and%20Accurate.pdf

Statistics Canada. "Statistics Canada's Quality Assurance Framework." 2002. http://www5.statcan.gc.ca/bsolc/olc-cel/olc-cel?lang=eng&catno=12-586-X

Viklund, Mattias. "Risk Policy: Trust, Risk Perception, and Attitudes." 2003. http://hhs.diva-portal.org/smash/get/diva2:221508/FULLTEXT01.pdf

Voida, Stephen, Elizabeth D. Mynatt, and W. Keith Edwards. "Re-framing the Desktop Interface around the Activities of Knowledge Work." *UIST '08: Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*. New York: ACM Press, 2008.

W3C Provenance Incubator Group. "Provenance XG Final Report." 2010. http://www.w3.org/2005/Incubator/prov/XGR-prov-20101214

Weinberger, David. "Transparency Is the New Objectivity." 2009. http://www.hyperorg.com/blogger/2009/07/19/transparency-is-the-new-objectivity/

Wf4ever Advanced Workflow Preservation Technologies for Enhanced Science. "Workflow Integrity and Authenticity Maintenance Initial Requirements." 2011. http://repo.wf4ever-project.org/dlibra/doczip?id=18

Wombacher, Andreas, and Mohammad Rezwanul Huq. "Towards Automatic Capturing of Manual Data Processing Provenance." 2011. http://doc.utwente.nl/77220/1/paper.pdf

Wootten, Ian, and Omer Rana. "Recording the Context of Action for Process Documentation." *Lecture Notes on Computer Science* 5272 (2008): 45–53.

Yeo, Geoffrey. "Representing the Act: Records and Speech Act Theory." *Journal of the Society of Archivists* 31, no. 2 (2010): 95–117.

Yeo, Geoffrey. "Bringing Things Together: Aggregate Records in a Digital Age." *Archivaria* 74 (2012): 43–91.