2017

Ethics in the Cloud

Corinne Rogers
University of British Columbia, School of Library, Archival and Information Science, corinne.rogers@ubc.ca

Luciana Duranti
School of Library, Archival and Information Studies, The University of British Columbia, luciana.duranti@ubc.ca

Follow this and additional works at: http://elischolar.library.yale.edu/jcas

Part of the Archival Science Commons

Recommended Citation
Available at: http://elischolar.library.yale.edu/jcas/vol4/iss2/2

This Article is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Journal of Contemporary Archival Studies by an authorized editor of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.
Ethics in the Cloud

Introduction

For the past several decades, information communication technologies (ICTs) have been changing the way we create, share, and keep our records and data. The handwritten letters and postcards that were the hallmark of personal written communication for centuries gave way to electronic bulletin boards and email, and these have been supplemented by text messages, tweets, and other forms of social media. Businesses moved from handwritten ledgers to electronic databases and shared drives. Today, individuals and organizations are increasingly creating, sharing, and storing information of all kinds in the cloud, with many of the same expectations of privacy, access, intellectual rights, and control that they have when storing it in in-house systems, either digital or analog. People are often surprised when they discover that behavior in the cloud is not guided by long-established ethical guidelines for information creation, sharing, and use. Instead, management of information and records in the cloud is controlled by legal contracts and enforced by laws, many of which are ill equipped to cope with the affordances of new technologies. Ethical expectations and guidelines that have been socially situated in a print culture developed over centuries are suddenly thrown into debate by technologies that are changing rapidly. What is the nature of information ethics in the digital era? In other words, is the expression “ethics in the cloud” an oxymoron?

In the context of the cloud, the ideas of privacy, access, intellectual rights, ownership, and control need to be reinterpreted and given new meaning. “An ethical society is based on a truthful understanding of what actually happened,” states Elena Danielson in her work, *The Ethical Archivist*.1 Today much of “what actually happened” has taken place online. But ethical considerations concerning presentation of information through traditional channels of communication do not translate seamlessly to online communities. This article explores the landscape of emerging ethical issues related to the creation, use, and maintenance of digital materials in cloud computing platforms in the course of our business and personal activities.

The Rise of Information Ethics

Three decades ago R. O. Mason argued that information forms the intellectual capital from which human beings craft their lives and secure dignity, and asked if the kind of society we are creating from this intellectual capital is the one we want.2 The question is even more relevant today. People’s intellectual capital is vulnerable in many ways. They may lose or reveal their personal information or find that it is used without their knowledge, permission, or compensation. They may be denied access to information that they should rightfully be able to see. Their personal information may be incorrect or they may receive information that is incorrect, either by mistake (misinformation) or design

---


Published by EliScholar – A Digital Platform for Scholarly Publishing at Yale, 2017
(disinformation). Based on these vulnerabilities, Mason identified four main ethical issues regarding information: privacy, accuracy, property or ownership, and accessibility.

Challenges in “information ethics” (IE) were raised as early as 1980, although ethical issues related to ownership and access long predate the digital era. IE initially focused on information and knowledge as resources, and ethical issues tended to concern the proper management of these resources, whether analog or digital. However, the affordances of digital written communication have shifted the focus from information as a resource to be managed and controlled, where the creator is the communicator, to information as a product to be used, where the human agent is the producer. IE is essentially concerned with who should have access to what information—core issues include intellectual freedom, equitable access, information privacy, and intellectual property. The shift of focus to use “also meant that ‘information ethics’ assumed greater urgency, as users and producers of information searched for fundamental rules that would prescribe how information should be responsibly collected, stored, and accessed.” According to Hauptman (the founder in 1992 of the Journal of Information Ethics), IE is fundamentally focused on the “production, dissemination, storage, retrieval, security, and application of information within an ethical context,” and addresses issues such as confidentiality, bias in information provided to clients or consumers, quality of data supplied by vendors, or use of work facilities.

Addressing Ethical Questions in Digital Information/Communication

Some scholars have advanced questions about whether IE was a new, autonomous field or a branch of applied ethics, and how one should justify actions based on ethical reasoning. IE may be considered a distinctive autonomous field in the sense that some acts involving computers present ethical qualities not possessed by any other type of act, but, while discussions are ongoing, a pragmatic approach suggests taking a traditionalist or conservative approach and regarding IE as a branch of applied ethics rather than a field dealing with radically new or exclusive moral issues. Regardless, justification for one’s actions requires ethical reasoning.

Professional fields are often guided by codes of ethics, sets of guiding principles for ethical behavior. Among records and archives professionals there is general consensus on core principles: to uphold intellectual freedom and resist censorship; to protect privacy

5 Fallis, “Information Ethics.”
6 Spinello, “Information and Computer Ethics.”
7 Ibid.
and confidentiality; and to recognize and respect intellectual property rights. These principles are aspirational, and therefore difficult to enforce.

Individuals often rely on popular heuristics to make ethical decisions. Many of these heuristics are blended systems of well-known and accepted rules. The Golden Rule states: “Do unto others as you would have them do to you.” The Platinum Rule states: “Treat others the way they want to be treated,” thereby shifting the focus of relationships from “this is what I want, so I will offer others the same treatment” to “let me understand what others want and then treat them accordingly.” Overriding both rules is the Do No Harm rule, even when others may ask you to act in a way that is harmful. Other rules consider the interplay of duties, rights, and responsibilities of interacting parties.

Traditional rules systems such as religious principles may guide ethical choices. Less formal heuristics include the “mom” test—which what would mom think; the “eye-team” test—what would the public think; or the “market” test—would you publicize your actions as a competitive customer relations strategy. Instead of applying moral rules, one may often resort to a “method of analogy”—comparing an unclear situation to one that is clear, drawing on the principle of consistency in reasoning (if X is right, and Y is essentially equivalent to X, then Y is right).9

Whether guided by an ethical code, heuristics, or analogy, one must still ask: what is the underlying ethical justification for an action? In addressing the issue of IE for library professionals, Fallis identifies four main types of ethical theories depending on whether they appeal to consequences, duties, virtues, or rights. For any ethical dilemma, one might ask what guidance would be given by a consequence-based, duties-based, rights-based, or virtues-based theory. Consequence-based theories support actions that have good consequences. Decisions arrived at through this lens have intuitive appeal and may be easily applied, but consequences are not the only things that matter, and many ethical dilemmas in dealing with ICTs arise out of the unintended and unforeseen consequences of actions. Duty-based theories hold that there are certain duties that must be upheld regardless of consequences. A few information ethicists espouse virtue-based theories that call for actions based on virtues such as courage, temperance, friendliness, or generosity. Several codes of conduct in online environments are based on these theories, but they are most relevant to communication between individuals rather than to the creation, use, and maintenance of information. Rights-based theories focus on actions determined by people’s rights by virtue of their status as human beings, or as members of society. These theories are the ones that most lend themselves to application to IE issues. One approach is that of John Rawls, who bases his theory on the idea of a hypothetical but fair agreement between people, used to evaluate large-scale social policies such as whether libraries should be publicly funded, or small-scale social policies such as how to establish reference and access protocols.10

**Information Ethics, Responsibility, and Trust**

---


10 Fallis, “Information Ethics.”
Regardless of the theoretical lens used to address ethical issues, two concepts are directly related to an understanding of IE: responsibility and trust. Ethical behavior of one person toward another implies a relationship of responsibility in which there is a subject, the entity or agent held responsible, and an object, the entity or agent to which the subject bears some responsibility. This relationship of responsibility needs to be supported by norms of behavior, and by the mechanisms that establish the relationship and maintain it so that it is workable. The norms are often prescribed by ethical considerations, for example, equitable access to information, and the mechanisms include contracts, laws, and other regulatory instruments. Relationships of responsibility that relate to ICTs include several significant issues that have led to the development of legislation and regulations. Two notable examples are privacy and data protection, and intellectual property.11

Research within the European Union on the ethical issues of emerging ICT applications (ETICA) and ongoing UK research on the “Framework for Responsible Research and Innovation in Information and Communication Technology” (FRIICT) have identified eleven technologies that are as likely to be socially and economically relevant in the coming ten to fifteen years as they are now; among them, cloud computing.12 Features shared by these technologies, all of which have ethical implications, include: natural interaction, invisibility of the technology, direct linkage through embedded or wearable tech, detailed understanding (by the technology) of the user (the human), pervasiveness, autonomy of ICT without direct user input, power over the user (i.e., the ability to structure the space of action of the user), and a market-driven focus.13

The responsibilities of different actors in these arenas are highly contested. Widely discussed ethical issues of emerging ICTs include privacy, security, trust, liabilities, and digital divides. It is interesting to consider whether predictions about emerging technologies can lead us to a better understanding of what could be done now in order to make sure that the ethical and social consequences of the technologies will be beneficial. New threats from emotion data and sentiment analysis (see, for example, recent suggestions that psychometric data may have been used in the 2016 US presidential elections).14


12 The eleven ICTs were Affective Computing—using computing to measure or express human emotions; Ambient Intelligence—ubiquitous and pervasive computing environment; Artificial Intelligence—representation of intelligence through artifacts; Bioelectronics—a combination of bio materials/principles with electronics; Cloud Computing—remote shared computing services; Future Internet—novel technical infrastructure for networked services; Human-Machine Symbiosis—direct combination of humans and machines; Neuroelectronics—the link between computing and neurosciences; Quantum Computing—the utilization of quantum effects for computing purposes; Robotics—embodied artificial agents, typically somewhat autonomous; and Virtual/Augmented Reality—the representation of reality through technical means.

13 Bernd, Eden, and Jirotka, “Responsible Research.”
elections\textsuperscript{14}) may raise qualitatively new issues. Less predictable ethical issues include human identity, the relationship between humans and technologies, and the relationships among individuals or groups.

If one expects to be treated ethically, there must also be a relationship of trust between the two parties. Traditionally, people’s and organizations’ trust in records and archives is based on four types of knowledge about their creator and/or their preserver: \textit{reputation}, which results from an evaluation of the trustee’s past actions and conduct; \textit{performance}, which is the relationship between the trustee’s present actions and the conduct required to fulfill his or her current responsibilities as specified by the truster; \textit{competence}, which consists of having the knowledge, skills, talents, and traits required to be able to perform a task to any given standard; and \textit{confidence}, which is an “assurance of expectation” of action and conduct the truster has in the trustee.\textsuperscript{15}

Trust may be defined as the confidence of one party in another, based on alignment of value systems with respect to specific actions or benefits, and involving a relationship of voluntary vulnerability, dependence, and reliance, based on risk assessment.\textsuperscript{16} To trust is to have confidence in another party with respect to specific actions or benefits. The four types of knowledge necessary for establishing trust are reflected in this definition of the trust relationship.

In the realm of cloud computing, one such relationship of trust is that between consumers of cloud services, as individuals or as communities of users, and cloud service providers (CSPs) in the consumption of cloud services. The mechanism through which trust is dictated is the service contract. Contracts for the provision of cloud services represent an important legal sine qua non for cloud use, and standardized provisions can avoid disputes and provide a fair balance between CSPs and users.\textsuperscript{17} However, the relationship between CSPs and users often reflects an imbalance of power: the user is dependent on the services of the provider with little or no chance of negotiating the terms of the relationship. While governments or large organizations have the capacity to negotiate the terms of their contract with these providers, most of us have no choice but to accept the boilerplate contracts written by the service provider. Boilerplate provisions are typically drafted by the dominant contractual party to suit its purposes, and are non-negotiable.\textsuperscript{18}

Few of us read these documents, and so we cannot be sure that our privacy will be


respected, our personal information not shared, or that our records will remain reliable and authentic.

The key question is to what extent a service contract is sufficient to establish trust in records from the perspective, knowledge, and requirements of a records manager or archivist. In order to embrace the contract as an instrument of trust, its terms must be transparent, understandable, and comprehensive for our needs. This demands that we articulate our needs and requirements at the outset with respect to issues such as data security, protection of personal information, availability of service, or location of data. Most of us would agree that all these are critically important, but we should also consider authenticity of records, demonstrable implementation of retention and dispositions schedules, and access to provider-generated metadata to prove provenance and chain of custody. These concerns, familiar to all records managers and archivists, are rarely considered priorities when we use cloud services, whether software as a service (SaaS), infrastructure as a service (IaaS), or platform as a service (PaaS).

**Main Issues of Information Ethics**

The four ethical issues of the information age that Mason identified in 1986: privacy, property rights, access, and accuracy, have become especially pressing in the cloud environment. Mason identified two great threats: the growth of IT (one might narrow that to ICT), and the increased value (and commodification) of information in decision-making.19

The privacy issues most relate to what information a person would be required to divulge about him/herself and about his/her associates, under what conditions, and with what safeguards. An invasion of privacy can be direct and immediate, or incremental, a condition that has been described as the “threat of exposure by minute description.”20 Massive amounts of data about individuals are held in the cloud and controlled by corporations and governments. Medical records, genotyping or gene sequencing data, medical history, prescription and insurance information, and test results and images are held by the medical establishment in private clouds, but still on the Internet, which makes them vulnerable to hacking. Genealogies, factual biographies, all sorts of biographic data and personal images are publicly available on the web, raising questions of ownership, copyright, and intellectual rights. Rights are further complicated in the case of a deceased person’s digital estate. Furthermore, the matching and integration of data has enormous ethical implications, in terms of power, potential for abuse, unauthorized access, error, and inaccuracy. Misinformation can wreak havoc in people’s lives, especially when the party with the inaccurate information has an advantage of power and authority.

Europe is developing a unified policy approach that will be difficult to harmonize with North America’s. In Europe privacy is considered a fundamental right and an aspect of dignity, while in North America it is an aspect of liberty and an alienable commodity that

---

19 Mason, “Four Ethical Issues.”
20 Ibid., 6.
can be renounced in order to have customized service or government protection. A frightening picture emerges when considering personal data. After 9/11 two things happened in parallel: private companies collected increasing amounts of personal data, and governments enacted laws demanding access to any and all data (new or old) held by corporations. These two factors have facilitated the gathering of huge amounts of personal data to benefit the business models of certain corporations, simultaneously enabling government surveillance. But the very protection of those private data makes it impossible for end users or society as a whole to reach conclusions about the authenticity or provenance of a particular claim, news story, or record: the information that could be used to trace them to their sources is protected by CSPs as private. The business model is based on an activity that is fundamentally predatory, since it repurposes other peoples’ productive content in a way that makes providers the maximum amount of money, and encourages unproductive, socially damaging activity, with no reference to the common good, however that is defined.

Contributing to the above situation, the technical infrastructures that gather and store data in cloud environments have become increasingly complex, hidden, and often invisible. Individuals have no idea which systems are collecting and sharing their data, or how to prevent them from doing so. Some groups and individuals feel a sense of injustice and are fighting back. They are protesting through the use of hacktivism, and they are using encryption and decentralized information processing technologies, such as the blockchain, to protect the “truth” and their privacy. However, we have no evidence yet that these technologies can be trusted.

Intellectual property rights (i.e., copyright and moral rights) are one of the most complex issues, with substantial ethical and economic concerns revolving around the attributes of digital information: easy to create, easy to share, copy, and disseminate once created, hard to safeguard once produced, hard to seek recompense when someone unauthorized uses it. Social media platforms facilitate the movement of material from one circle of people to another and digital information is easy to repurpose and reuse. Reuse, however, is often remix, a practice that results in derivative works that substantively change the intent and context of the appropriated material. Ad hoc dynamic groups collectively create a body of interlinked material related to a common interest further complicating understanding of ownership and provenance. Social norms are beginning to emerge through successive cycles of use, reuse, modification, repurposing, and take-down notices, but often, people would rather ask for forgiveness than for permission.

22 Cullen Hoback, Terms and Conditions May Apply, News Documentary, 2013.
26 See https://syrianarchive.org.
Access to digital material, related to availability, presents ethical questions in the cloud environment. Where access to information is a right, availability should be a fact, but the former cannot be satisfied without the latter. Despite sometimes significant jurisdictional differences, legislation in North America and elsewhere exists that guarantees the right to certain information held by various public bodies, and sometimes also by private organizations. In some jurisdictions this information must be provided within a specific period of time. When the data are stored in a cloud environment, “availability of the stored data implies also the availability of the infrastructure, hardware and software, which facilitates the retrieval and readability of the data,” because technical difficulties might slow the process, and the owner of the data, being liable for providing access to them, may be sanctioned.27 Furthermore, availability, that is, “the amount of time that a system is expected to be in service,” expressed either statistically or as a percentage, is linked to “reliability,” the characteristic of behaving consistent within expectations.28 Thus, one must consider not only availability but also “consistency and accuracy of access.” This means that copies of the data must be distributed across several data centers, ensuring redundancy, but also that such copies must remain consistent while users access the same data at the same time. This is not currently possible as providers do not have explicit agreements with each other that help ensure the reliability of the Internet overall. The latter will require collaboration among multiple regulatory authorities, service providers, users, security/public safety communities, and international trade and standardization communities.29

Access issues are also linked to information literacy. Literacy is a requirement for full participation in society and each innovation in information handling places new demands on literacy. Access requires intellectual skills to deal with information (reading, writing, reasoning, calculating) as well as the ability to deal with the information technologies that store, convey, and process information, both responsibilities of the education system. Also, we know that what people read on social media is filtered: the stories reinforce members’ beliefs and those of their friends. After all, these stories are selected by algorithms that make the most money for corporations like Facebook, which operates a network of 79 percent of online American users.30 “On Facebook, what you click on, what you share with your ‘friends’ shapes your profile, preferences, affinities, political opinions and your vision of the world. The last thing Facebook wants is to contradict you in any way. The sanction would be immediate: you’d click/share much less; even worse, you might cut your session short. Therefore, Facebook has no choice but keeping you in the warm comfort of the cozy environment you created click after click. In the United

27 Bushey, Demoulin, and McLelland, “Cloud Service Contracts.”
29 Ibid., 100–101.
States, Facebook does this for 40 minutes per user and per day.”31 The result is a sort of dystopian social realization of Thomas Kuhn’s observation that the nature of scientific fact is not solely based on objective criteria but shaped by the opinions of a community.32

Access is also linked to ownership. For example, an individual’s digital fond, or archive, is a mix of what is purposefully kept and what is forgotten in cyberspace (one’s digital shadow), as well as the residue of Do It Yourself (DIY) archiving (including other entities’ and people’s records). Ownership of this material does not always reside with that individual. When an individual dies, ownership of and access to their records is further complicated. Service providers share ownership or use of the data, and may or may not let the next of kin download copies in useful forms.

With so many ethical issues mounting, how can the records and archival professional deal with the cloud environment without feeling like s/he is moving in the dark, subject to unpredictable consequences? Blanchette states that the cloud has become a “certain kind of meta-infrastructure” capable of unprecedented sustainable growth, where infrastructure is “defined as the elements of the computing ecosystem that provide services to applications, in contrast to the applications that provide services to users.”33 This means that countries must begin to look at the cloud as a critical infrastructure, that is, one that is vital to the functioning of their economy and society. The fact that public recordkeeping and archival preservation are increasingly entrusted to the public cloud would both support such a determination on the part of governments and facilitate the choice of the public cloud for the records and archives of businesses and non-public organizations. However, critical infrastructures are dependent on other infrastructures and some cloud services depend not only on electrical and communication infrastructures, but also on other cloud services. Blumenthal believes that there is potential for federated clouds to assist one another by sharing resources in the event of a crisis.34 Many have been calling for an international, cohesive framework of policies with regard to the cloud environment. Among them, the European Commission has been the most active. At its 2015 cloud security conference, it was agreed that there is a need for both flexible policy approaches allowing for technological advancement and a stronger relationship between the public sector and private industry, establishing security in terms of networks, data location requirements, foreign jurisdiction, and access.35 There is no reason why such approaches should not be extended to ethical issues.

The cloud is the platform of choice for mobile applications and the data generated using them, as well as those created in smart devices at home and at work, and records creators generate a growing percentage of data in the public cloud and have to rely on them for their keeping and preservation; hence, the existence of such a federation would facilitate the creation of a shared ethics code. At least this is our hope for the future to ensure that ethics in the cloud will not be an oxymoron.

Bibliography


