Science Through the "Golden Security Triangle": Information Security and Data Journeys in Data-intensive Biomedicine

Research-in-Progress

Niccolò Tempini
University of Exeter
Byrne House, St German's Rd, Exeter, EX4 6TJ, UK
n.tempini@exeter.ac.uk

Abstract

This paper talks about ways in which infrastructure for biomedical data-intensive discovery is operationalized. Specifically, it is interested in information security solutions and how the processes of scientific research through data-intensive infrastructures are shaped by them. The implications of information security for big data biomedical research have not been discussed in depth by the extant IS literature. Yet, information security might exert a strong influence on the processes and outcomes of data sharing efforts. In this research-in-progress paper I present a developing, in-depth study of a leading information linkage infrastructure that is representative of the kind of opportunities that big data technologies are occasioning in the medical field. This research calls for IS to extend the discussion to consider, building on the empirical detail of intensive case studies, a whole range of relations between provisions for information security and the processes of scientific research and data work.

Keywords: Information security/privacy, Information infrastructure, Data resource management, Health informatics/health information systems/medical IS, Information, Information management, IS security, Big data, Data-intensive Science, Data security

Introduction

This paper talks about ways in which infrastructure for biomedical data-intensive 'discovery' is operationalized in big data research infrastructures. Specifically, it is interested in information security solutions and how the processes of scientific research through research infrastructures are shaped by them.

The implications of information security for data-intensive biomedical research have not been discussed in depth by the extant IS literature. Discussions of the consequences of norms and regulations for accessing data in medical journals and reports have mostly been concerned with highlighting that while protection of individuals must be guaranteed, complex ethics review processes make research more difficult (Dove et al. 2016) – as researchers need to involve in lengthy negotiations for accessing less data than they ideally would. This research includes these concerns but calls for IS to extend the discussion to consider, building on the empirical detail of intensive case studies, a whole range of relations between provisions for information security and the processes of scientific research and data work.

This must not be narrowly understood as a range of relations that is contained within the boundaries of a single organization, as in the split between production and support functions that characterizes much of the literature on security and work process. Not only the operational boundaries but also the institutional reference points of an organization have become blurred as a result of the great reconfiguring of relations

that networking technologies and data-intensive work have occasioned. On this backdrop, this paper supports the view that in order to develop systems that allow data-intensive biomedical research to happen, technical and governance controls must be continuously related with external stakeholders representative of the wider society here involved (Burton et al. 2015).

The research-in-progress I am presenting here intends to address a gap in our understanding of the implications of information security by trying to link information security requirements and solutions with the shape that opportunities for medical discovery take in new, cutting edge data sharing infrastructures. I present a developing, in-depth study of a leading *information linkage infrastructure* that is representative of the kind of opportunities that big data technologies are occasioning in the medical field. This kind of infrastructures collect data that have been generated by a number of institutions in a specific territory as routine task in the context of health, social and mental care service delivery (these include but are not limited to, for instance, data from GP –general practitioner– practices or hospitals). These datasets are linked at individual level, anonymized, in order to construct the possibility of longitudinal studies that follow individuals between institutions and facilities, over time, and throughout the course of their health life.

In the research I intend to show how, in different ways across different contexts, the flows and life-cycles of data and the directions in which research process consequently develops are shaped by governance panels, access rules and enforcing technologies. The argument is of interest to the information systems community, which holds a deontological commitment in linking practices and technologies to their societal and organizational consequences. It contributes to an understanding of the conditions that make data-intensive science through big data technology possible or impossible.

In this paper, I argue that the role of information security in data-intensive science has been left undetected. Yet, I argue, information security might exert a strong influence on the processes and outcomes of data sharing efforts. Expectations on the re-use of scientific data, as raised by arguments highlighting the opportunities of open data and big data infrastructures, and reified by the funding that is dedicated to these initiatives, rely on the hope of multiplying scientific investigations with the re-used data. By and large the idea is that purposefully developed infrastructures are important to allow multiple actors to look at datasets with new eyes. The opportunity to explore data is deemed important for generating new insights.

However, socio-technical environments engineered to control and govern the access to these resources through a mixture of response and preventive technologies (Baskerville et al. 2014) and procedures could well shape the trajectory of research processes and projects. And yet, security solutions and architectures are related to a fast-changing environment through the hazards that it presents (Baskerville et al. 2014), but also to the requirements that it formulates. Information security practice is embedded in the material, organizational, social and regulatory landscapes of research: requirements are deeply context-dependent. Data infrastructures for biomedical research in particular face very tight requirements on data handling, storage, access and re-use.

Privacy is paramount in medical research and the embedment of its definition at the heart of society sometimes has been ignored by researchers and infrastructure developers alike – at their own peril. The debacle of the "care.data" initiative and infrastructure in the UK is one of the most recent and a very illustrious example of how a technical infrastructure fails when its definition, discussion and development happen in neglect of its relations to the wide and complex of society (Ward 2015). For a long time, since well before the development of the data-intensive infrastructures here discussed in this paper, the protection of patient privacy and questions of public good have been an ineludible pre-condition for research in biomedicine involving human subjects (directly or indirectly). Yet and importantly, the mutable definitions of these continuously constructed concepts necessitate an exercise in reflexivity to be situated at the heart of each infrastructure project, to consider emerging risks, and which needs to be repeated in the face of changing environmental conditions (Ciborra 2007).

How information security requirements are defined depends among other factors on the disciplinary, legal, ethical, political and business context in which the infrastructure is situated. Developing infrastructures in a sensitive, contended and competitive ground like that of biomedical research demands the implementation of complex solutions. Whether data consigned to databases are accessible and useable depends partly on the strategies employed by database developers and curators to keep data

alive as potential evidence and valuable commodities. The data that these infrastructures host must be maintained and made available in ways that are useful for researchers and potentially productive of new knowledge, while at the same time the infrastructure must protect itself and the stakeholders from unwanted consequences (for instance unauthorized disclosure, and individual re-identification).

On this backdrop, I argue that there is a need for understanding about what these requirements mean in the development of data-intensive research systems and their eventual implications for science practice. For instance, information security solutions aiming at protecting patient confidentiality can require information suppression at several stages of the data journey process: how could this shape the trajectory of a research project and its end outcomes? These are the kinds of questions that I set out to investigate in this research project. To investigate the implications of information security on the organization of the research process (Lowry et al. 2015), we must document the pipeline of data sharing, collecting information about all the different steps involved in making re-use of data in research possible (Leonelli 2014). As Nissenbaum argued, a fine-grained study of systems is fundamental to understand the social dimensions of new technology (Nissenbaum 2001).

Information Security and Data-Intensive Work

Literature on information security in information systems has not yet been concerned with the question of the role of information security functions in big data or data-intensive science infrastructures. Still, the literature employs a number of concepts that can be helpful as analytical lens to guide this investigation, however, it is not straightforward how to inform this case. Periodically, reviews of the extant literature have lamented how the security field has remained theoretically underdeveloped, mostly concerned with sharing anecdotal evidence and recommendations in which social and management issues are discussed in negligible proportions (Siponen et al. 2008). Most security development methods have long been largely unable to thoroughly factor in the unique social and human factors that characterize every organization and its work processes (Siponen 2005). Unfortunately, translating insight from anecdotal recommendations to new analytical work is difficult when the phenomenon of investigation is an innovative and complex infrastructural arrangement.

Despite these critiques, the one conceptual framework that has remarkably shaped practice and research is the "golden security triangle" of CIA (Shameli-Sendi et al. 2016:23) that divides risks of attacks to data assets between, namely, Confidentiality, Integrity and Availability (Gold 2010; Nissenbaum 2005).[1] Security incidents affect confidentiality when protected information is disclosed to unauthorized parties, integrity when data and software are corrupted and original information is compromised, availability when information is made inaccessible due to system shut down or service disablement.

Information security risk assessment approaches (Shameli-Sendi et al. 2016) help us to understand that health datasets held by an information link data infrastructure are critical targets, which can be targeted by attacks, with potential long-term consequences. Surely, linked health data are valuable because they can be linked to other data to produce new information. Also, they are sensitive because they can be used for the re-identification of vulnerable and unaware individuals. As a result, the legal, social and normative pressure on health data infrastructures is extremely high. Here, risk means more than losing control of a valuable asset after data breach and disclosure. Risk of a security accident can impact the organization at different levels of abstraction (asset, service, and business), and when there are dependencies with other resources at the same or different levels, risk propagates in complex ways (Shameli-Sendi et al. 2016), potentially endangering the enterprise.

Once data are held, they can be put to use in multiple ways and for a potentially long time. The social and organizational costs of a consistent data breach of linked health datasets could be very high. As Graves and colleagues highlight (Graves et al. 2016), costs to be accounted for a data breach are not only direct ones but also the indirect ones. In a data infrastructure for biomedical data linkage research, indirect costs could include to make future data sharing more difficult both for the affected infrastructure and for other infrastructures. The existence of large stolen datasets would increase the risk of linkage with other datasets at the attacker's disposal, increasing the incentive for further attacks in turn. The impact on

[1] It must be said that there have been proposals for modifications and extensions to the CIA triad of categories, but it remains out of the scope of this paper to discuss them.

public confidence could mean tighter regulations, and fewer participation in research studies due to increased privacy concerns. As Durante explains (Durante 2011), the definition of privacy is not about the control over specific data assets, but rather in the right to participate in the construction of one's life self-narrative. As such the potential damage to the right to privacy from re-identification of linked, longitudinal health data profiles is extremely high, as linked datasets are multi-dimensional and comprehensive.

To contribute to the uncertainty as to what measures are appropriate to the risks involved in developing a data sharing infrastructure, the ethical challenges implied by data sharing and linkage are the object of an open discussion.[2] Several years after Nissenbaum's argument on the convergence of computer and national security (Nissenbaum 2005), there is little clarity as to how we should consider risks involving publicly-funded research infrastructures (Stahl et al. 2012), which one might argue sit both at the critical infrastructure level and at the level of the single organization concerned with protecting from attacks and related (consequential) liabilities.

In addition, the NHS in UK is still a main target of security breaches, with implementation of information security policies an open problem that can be affected by both insufficient level of guidance by management concerned with discharging responsibilities (Stahl et al. 2012), and lack of care and skill by employees that operate data in NHS institutions (Gold 2010). All these circumstances suggest that mapping the ramifications of risk is a very complex and uncertain undertaking and might be mobilized in favor of highly-protective approaches to security of an infrastructure, which could come to rely on an array of "hard trust" technical solutions at the expense of "soft trust" based on social relations (Patrick et al. 2005).

Flipping the "Golden Triangle of Security"

According to Stahl and colleagues, we do not know enough on the organizational role of information security policies and arrangements (Stahl et al. 2012). As such, it becomes interesting to ask how does information security in a high-risk setting such as a health data infrastructure impact its core function, that of enabling scientific research. Scientific research can easily be understood as a kind of work process that is difficult to pin down to a model and a list of requirements.

A main concern for researchers of usable security is that security should not involve trade-offs that negatively shape work processes in the organization. When security makes work more complex, people prioritize execution over security. Designing systems that are psychologically acceptable requires designing for the expectations of those who will use the system (Bishop 2005; Sasse and Flechais 2005). This in a data infrastructure might be particularly difficult since end users are health researchers, who are often distributed both geographically and institutionally. Most often they have no training in using security technology, and it can be easy for developers to assume sophisticated user understanding of information sharing (Good and Krekelberg 2003). Despite health researchers are consistently reminded of the risks involved in handling human subject data, lack of knowledge about users makes it difficult to trust them and design the most usable technology (Adams and Sasse 1999). Also, security is usually understood as a supporting task (Sasse and Flechais 2005) and not a production task, and this separation might perhaps contribute to make it difficult to consider its role in shaping data-work processes, as it is not seen as a defining element of work. Unmotivated users, abstract policies and lack of user feedback

[2] For instance, Aicardi and colleagues argue that "changing practices in the collection and use of digital data require a revised framework and nomenclature regarding the norms, rules, and principles governing biomedical research" (Aicardi et al. 2016:209-210) and indicate five open challenges: 1) the concept of "personal data" is changing, as data about an individual can be personal for more than one person; 2) it is not possible to guarantee definitive anonymization in practice; 3) it is impossible to anticipate all future uses of the data, therefore informed consent is at issue; 4) a much broader spectrum of datasets can be put to health-relevant use (think search engine queries) than those collected by accordingly regulated institutions; 5) there are limited options to rectify consequences of inferencing errors at a time when predictive analytics are gaining momentum. As the authors suggest, these challenges are open to discussion and are an opportunity for all affected parties to make a contribution towards solutions.

united with the criticality of potential breaches and the vulnerability of any network from its weakest link are all factors that conjure up against design of technology with ideal usability (Whitten and Tygar 1999).

While the usability literature helps to understand how security can be perceived as an obstacle, I am interested in understanding in what ways *compliance* can shape the supported research work process. If a support function interferes with a production function in an infrastructure, this is important. This is particularly sensitive in data-intensive work such as research in big data infrastructures, I argue, as work processes are deeply involved in manipulating the very same resource that information security requirements might need to be as controlled as possible. The distinction between supporting and production task, in this light, tends to blur. Designing enabling information security becomes of paramount importance for research infrastructures that are aimed at maximizing re-use of the data. Even when implemented information security can be unquestionable given legal, social and ethical circumstances, it is important to understand if it has implications. Still, understanding the costs of information security is not an easy task also because it is difficult to estimate the opportunity cost of missed research partnerships. But I have tried to argue, to this point, that it is still worth asking the question.

On this backdrop, in this research I try to operate a "gestalt switch" and "flip" the CIA framework to include not only external risk factors such as attacks and theft, but also and most importantly internal ones, including information security itself and its solutions. Processes that are internal to the secure infrastructure might stand in an ambivalent relationship with the research process itself and its potential to detect and track the phenomena of interest –scientific research is very susceptible of the specific material arrangements in which it is carried out.

Methodology

The research builds on an ethnographic study of a data linkage infrastructure in the biomedical research domain that is based in the UK: the Secure Anonymised Information Link (SAIL). SAIL is a databank developed to make possible the re-use for health research purposes both of routine data generated through public services and of otherwise unavailable datasets generated in the context of individual scientific projects. The publicly-funded infrastructure was developed within the Health Informatics Research Unit at the University of Swansea in Wales to build a world-class health research facility for research communities both locally and globally, aiming at re-purposing biomedical datasets through their integration in a dedicated digital infrastructure. Given the amount of sensitive information involved in such an enterprise, SAIL needs to address concerns around data privacy, and both the public good and potential harm that such an infrastructure can generate.

I have argued in the introduction that in order to be able to study the data re-use pipeline it is necessary to collect information about all the steps involved in the "data journey" (Leonelli 2016; Leonelli 2014). The focus of the research is not "simply" to situate a data-intensive research infrastructure in the wider context and discuss how the environment offers opportunities and exercises pressures on the infrastructure, as from an external point of view. Instead, it is about tracing in detail how these conditions shape the data management and research processes (and outcomes) that the infrastructure aims to enable. As such, current conceptualizations of data re-use journeys such as, for instance, digital data streams (Piccoli and Pigni 2013) are not well developed in this respect. The focus should not be whether an infrastructure is an aggregator, a creator, or else, of data. Rather, it should be to look at all the steps of data work and understand their implications for final outcomes (Tempini 2015). It is about a different way to attend to value creation, at the level of actual data practices. This level of detail is crucial to be able to make empirically-informed claims about scientific practice and its knowledge creation regimes.

To this end, I have conducted more than 20 interviews with relevant SAIL members at all levels, covering scientific, technical and support development tasks. Crucially, I have also interviewed researchers external to the organization who have used SAIL for research purposes, i.e. the end users. Interviews are semi-structured, following an itemized interview-guide that is custom to every interviewee (Flick 2006). Data collection includes as well internal documentation, presentations, and observations collected in SAIL analyst meetings. The topic of information security and its impact in the research process emerged very early in the interview process and was since integrated as a main topic. The data collection phase is continuing. Only a small portion of interviews has so far been analyzed with a first coding pass, to review

progress, identify emerging themes for analysis, and select topics for further investigation in interviews. What follows is not a complete analysis of the data thus collected. Instead, I offer in the form of a short empirical narrative a review of the areas of investigation that I have been concerned with in respect to information security and research processes.

First observations: governing inferences, shaping information flows

SAIL is a databank developed to make possible the re-use for research purposes of routine data generated through public services but also datasets generated in the context of a scientific project. Here, researchers need to satisfy demanding criteria in order to access and use the data, and publish the findings. Monitoring and reviewing processes bring the research project to run along a well-established path. Research questions are discussed and reviewed before the application and the actual start of a research project. At the same time, the staff operating the infrastructure needs to satisfy also other demanding criteria to guarantee the security of the research environment, and technology is deployed in order to constrain human agency and ensure compliant behavior.

A considerable and, for research purposes very important, proportion of the data handled by SAIL is collected by GP practices, where practitioners are invested with the duty of managing and protecting patient records. When discussing their role in transferring such data to a centralized infrastructure, database managers describe their duties as "data custodian" or "guardian", with responsibilities to guarantee the security of the research environment. In fulfilling the role of data guardians, SAIL staff are faced with a variety of requirements that need to be satisfied for the infrastructure to attract researchers and dispel risks and polemics.

Information security concerns are paramount and determine the specific solutions adopted to safeguard data access and use. In a stark contrast with bird-eye views (and claims) for big data research to have no constraints of scope nor scale, these concerns continuously shape choices about the tools, the time and the data that are provided to researchers in the context of each project. To this aim research activity with SAIL data is confined to a carefully designed virtual research environment: access is strictly regulated (with automated access restriction to specific users, at specific times of access, and to specific data resources), user behavior is automatically monitored and can be audited at any time, and functionality which constitutes an unacceptable information security risk (most importantly, the unauthorized exportation of data) is restricted. All data are anonymized through a complex split-file approach, to guarantee that no party holds access to the entirety of the linked information.

In combination with cutting-edge information technology solutions, security of the research projects is ensured through a (more traditional) review process of project applications involving external stakeholders, and the following negotiation of an economic set of data variables that is deemed to be sufficient to carry out the project and which will be made available in the research environment. In line with other databases storing sensitive datasets, the review panel members need to investigate and evaluate the motivations, methodologies and aims of their users against the crucial dimensions of feasibility and risk. Applicant researchers, review panel members and SAIL analysts are routinely involved in a step-wise triangulation process where the understanding on one hand of what is feasible from a research point of view, given the available data sources, and on the other hand of what is sensible and secure research, is constructed each time through a highly contextual process. Information security requirements are deeply situated and localized in each project, and defined as much by the nature of the data types and phenomena in question, as by the disciplinary, legal, ethical, political and business context in which the infrastructure is situated.

Crucially, these security arrangements are intended to shape information flows, to selectively govern the possibilities for inferencing about individual health records. On the one hand much security is dictated by the need to reduce to the minimum the risk of deductive disclosure (the risk of disclosing protected information –i.e. individual identity– as a result of linking together datasets). One the other hand, the infrastructure is shaped and designed to enable as much as possible very different kinds of inferences from the data, such as those investigating relationships within individual health histories.

Importantly, at the aggregate level decisions about individual projects, and the experience of the collaborations that further unfold, can have deep effects, reaching far beyond the individual project. SAIL is a full-fledged research infrastructure that has evolved over time through negotiations that interface the

requests advanced by researchers seeking to create new knowledge and the concerns, conventions and regulations that require the research projects to be transparent and accountable. The complex solutions that have evolved over time are the result of this repeated interfacing. They are continuously re-assessed and modified for the improvement of one or another aspect, and are at the same time the basis for incremental improvement and path-dependent infrastructural inertia.

Concluding observations: for the security of security solutions

Achieving information security is broadly defined as a matter of triangulating accessibility (usable and practical access to resources), confidentiality (protection of sensitive information) and integrity (data fidelity is ensured) against the social context in which an infrastructure is situated. When analyzing the SAIL material, it appears that information security is not only concerned with risk from unauthorized action from 3rd parties. In consideration of the sources of potential loss of productive function, this research opens up the question as to whether requirements towards three CIA dimensions sometimes work at cross-purposes – for instance, we might ask if confidentiality provisions might reduce themselves the accessibility of data assets. If datasets are an asset that must be protected, to be employed in services for data re-use, then information security investigations can concern processes *both* inside vs outside the organization as well as compliant vs non-compliant. If we define risk as "the exposure to any proposition which involves uncertainty" (Holton in Shameli-Sendi et al. 2016:16), risk might be inherent to certain kinds of operations with data that pertain to secure data management and infrastructure development. Eventual operations of data management that might turn a dataset unable to be put to work for the production of new information become of concern under this extended analytical lens. A certain form of risk recursivity seems to be possible: that an information security solution at an asset level might become a source of risk for running a service or the business level. Source of risk interact if they require countermeasures that work at cross-purposes.

In consideration of the evidence thus far collected and reflected upon, I believe it is necessary to further open up a discussion about how information security requirements and solutions shape the trajectories and outcomes of efforts for data sharing for research. This is a momentous opportunity for IS scholarship to make a broad contribution as, in the time of Big Data, initiatives are multiplying that seek to link health datasets with one another and with datasets originated in diverse settings such as administrative or the natural sciences. IS scholarship needs to engage to understand the conditions that shape the journeys of data.

Acknowledgements

This research is funded by the European Research Council under the European Union's 7th Framework Programme (FP7/2007-2013) / ERC grant agreement n° 335925.

References

Adams, A., and Sasse, M. A. 1999. "Users Are Not the Enemy," *Communications of the ACM* (42:12).

Aicardi, C., Del Savio, L., Dove, E. S., Lucivero, F., Tempini, N., and Prainsack, B. 2016. "Emerging ethical issues regarding digital health data. On the World Medical Association Draft Declaration on Ethical Considerations Regarding Health Databases and Biobanks," *Croatian Medical Journal* (57:2), pp. 207–213.

Baskerville, R., Spagnoletti, P., and Kim, J. 2014. "Incident-centered information security: Managing a strategic balance between prevention and response," *Information & Management* (51:1), pp. 138–151.

Bishop, M. 2005. "Psychological Acceptability Revisited," in *Security and Usability. Designing Secure Systems That People Can Use,* L. F. Cranor and S. Garfinkel (eds.), Sebastopol, CA: O'Reilly Media, pp. 1–12.

Burton, P. R., Murtagh, M. J., Boyd, A., Williams, J. B., Dove, E. S., Wallace, S. E., Tassé, A.-M., Little, J., Chisholm, R. L., Gaye, A., Hveem, K., Brookes, A. J., Goodwin, P., Fistein, J., Bobrow, M., and Knoppers, B. M. 2015. "Data Safe Havens in health research and healthcare," *Bioinformatics* (31:20), pp. 3241–3248.

Ciborra, C. 2007. "Digital technologies and risk: a critical review," in *Risk, complexity and ICT*O, Hanseth and C. Ciborra (eds.), Northampton, MA: E. Elgar, pp. 23–45.

Dove, E. S., Townend, D., Meslin, E. M., Bobrow, M., Littler, K., Nicol, D., Vries, J. de, Junker, A., Garattini, C., Bovenberg, J., Shabani, M., Lévesque, E., and Knoppers, B. M. 2016. "Ethics review for international data-intensive research," *Science* (351:6280), pp. 1399–1400.

Durante, M. 2011. "The Online Construction of Personal Identity Through Trust and Privacy," *Information* (2:4), pp. 594–620.

Flick, U. 2006. *An Introduction to Qualitative Research*, Thousand Oaks, CA: Sage.

Gold, S. 2010. "Securing the National Health Service," *Computer Fraud & Security* (2010:5), pp. 11–14.

Good, N., and Krekelberg, A. 2003. "Privacy and Trust, Usability and Privacy: A Study of KaZaA P2P File Sharing," in *Proceedings of the Conference on Human Factors in Computing Systems (CHI '03)*, Ft Lauerdale, April 5.

Graves, J. T., Acquisti, A., and Christin, N. 2016. "Big Data and Bad Data: On the Sensitivity of Security Policy to Imperfect Information," *U. Chi. L. Rev.* (83), pp. 117–457.

Leonelli, S. 2014. "What difference does quantity make? On the epistemology of Big Data in biology," *Big Data & Society* (1:1), (doi: 10.1177/2053951714534395).

Leonelli, S. 2016. *Researching Life in the Digital Age: a philosophical study of data-centric biology*, Chicago, IL: University of Chicago Press.

Lowry, P. B., Dinev, T., and Willison, R. 2015. "Call for Papers: European Journal of Information Systems (EJIS) Special Issue on Security and Privacy in 21 st Century Organisations," *European Journal of Information Systems* (available at http://perma.cc/53DS-TEMS).

Nissenbaum, H. 2001. "How computer systems embody values," *Computer* (34:3), pp. 120–119.

Nissenbaum, H. 2005. "Where Computer Security Meets National Security," *Ethics and Information Technology* (7:2), pp. 61–73.

Patrick, M. A., Briggs, P., and Marsh, S. 2005. "Designing Systems That People Will Trust," in *Security and Usability. Designing Secure Systems That People Can Use,* L. F. Cranor and S. Garfinkel (eds.), Sebastopol, CA: O'Reilly Media, pp. 75–100.

Piccoli, G., and Pigni, F. 2013. "Harvesting External Data: The Potential of Digital Data Streams," *MIS Quarterly Executive* (12:1), pp. 143–154.

Sasse, M. A., and Flechais, I. 2005. "Usable Security," in *Security and Usability. Designing Secure Systems That People Can Use,* L. F. Cranor and S. Garfinkel (eds.), Sebastopol, CA: O'Reilly Media, pp. 13–30.

Shameli-Sendi, A., Aghababaei-Barzegar, R., and Cheriet, M. 2016. "Taxonomy of information security risk assessment (ISRA)," *Computers & Security* (57), pp. 14–30.

Siponen, M. T. 2005. "Analysis of modern IS security development approaches: towards the next generation of social and adaptable ISS methods," *Information and Organization* (15:4), pp. 339–375.

Siponen, M., Willison, R., and Baskerville, R. 2008. "Power and Practice in Information Systems Security Research," in *ICIS 2008 Proceedings*, Presented at the International Conference for Information Systems, Paris, France, December 14-17 (available at http://aisel.aisnet.org/icis2008/26).

Stahl, B. C., Doherty, N. F., and Shaw, M. 2012. "Information security policies in the UK healthcare sector: a critical evaluation," *Information Systems Journal* (22:1), pp. 77–94.

Tempini, N. 2015. "Governing PatientsLikeMe: information production and research through an open, distributed and data-based social media network," *The Information Society* (31:2), pp. 193–211.

Ward, A. 2015. "Health advances using 'big data' at risk, ministers warned," *Financial Times* (available at http://www.ft.com/cms/s/0/870ca844-aaf3-11e4-81bc-00144feab7de.html?siteedition=uk).

Whitten, A., and Tygar, J. D. 1999. "Why Johnny Can't Encrypt," in *Proceedings of the 8th USENIX Security Symposium*, Washington, DC, August 23, pp. 169–184.