Privacy by Obfuscation with Personal Data Management Architectures: Possibilities and Constraints

Dave Murray-Rust
University of Edinburgh
School of Informatics
d.murray-rust@ed.ac.uk

Kieron O’Hara
University of Southampton
Electronics and Computer Science
Highfield
kmo@ecs.soton.ac.uk

Marion Oswald
University of Winchester
Centre for Information Rights
Marion.Oswald@winchester.ac.uk

Max Van Kleek
University of Southampton
Electronics and Computer Science
Highfield
emax@ecs.soton.ac.uk

Nigel Shadbolt
University of Southampton
Electronics and Computer Science
Highfield
nrs@ecs.soton.ac.uk

ABSTRACT

In this position paper, we discuss legal and technical aspects of protecting privacy using Personal Data Management Architectures (PDMAs), which include, but are not limited to Personal Data Stories and Personal Information Management Services. We argue that providing false information on occasion is a common strategy online and offline for people to protect their privacy and determine their representation in the world, and we discuss some empirical findings to that effect. We describe a potential, and technically-feasible, ecosystem of digital practices and technologies to facilitate this practice, and consider what legal frameworks would be required to support it.

Categories and Subject Descriptors

K.4.1 [**Computers and Society**]: Public policy issues – *privacy, use/abuse of power.* K.5.m [**Legal Aspects of Computing**]: Miscellaneous – *contracts.*

General Terms

Security, Human Factors, Legal Aspects.

Keywords

Personal data stores, personal information management, obfuscation, anonymisation, data protection, privacy, lying.

1. INTRODUCTION

Online interaction threatens to reduce our ability to create and curate plural identities. We are used to having awareness of and control over the ways in which we are perceived by others in the multitude of social situations we navigate on a daily basis, and it is understand that, for example, one might be a different person in the bedroom, behind closed doors, to the professional persona that one adopts at the workplace [1].

Online actions become data, which is not strongly bound to the spatiotemporal context in which it was collected. boyd warns of *context collapse* – the aggregation of the diverse facets of personhood into a single overall identity as actions taken in one context spill out into other places [2]. Lyon uses the metaphor of *leaky containers* – information held in one place leaks, corroding compartmentalization, seeping into places where social norms and conventions are different [3]. Butler, following Althusser, uses the term *interpellation* to describe the ascription of identity to people by others [4]. The application of labels and their acceptance gives power to the labeler [5]. Both the scope of the data being collected and the inferences made from it are open to expansion without knowledge or control of its subjects. As such, they are interpellated without consent or understanding into categories and roles devised by others, of unknown scope, influence and intention.

Participation in online socialization is increasingly prevalent; resistance and refusal rare and isolating. How then, can we regain control over labeling and sorting, and what do we have to trade in order to do this? Our aim in this position paper is to explore one possible solution, the use of technical means to manage the exploitation of personal data. This is a complex area, and our focus is the regulatory background to such architectures concerning the data subject’s control over the information given out and its veracity. In the offline world, such control is an important aspect of informational self-determination. Based on the principle of contextual integrity introduced by Nissenbaum [6], it could be argued that this control should be re-established online, via technical or regulatory means. In this paper, we begin with a discussion of people’s manipulation of information to preserve aspects of identity, and then consider how technology could support these processes. We then consider the legal context which may apply to such technologies.

2. ONLINE LYING

It is broadly accepted as a moral dogma that lying is wrong. Deontic philosophers such as Augustine, Aquinas and Kant oppose lying *tout court*. Consequentialists accept a notion of a ‘white lie’, where the outcome of truthtelling would be worse than the opposite, but most would argue that these situations are the

Conference’10, Month 1–2, 2010, City, State, Country.

exception. The ideas of communicative rationality promoted by people like Habermas require a discourse ethics that assumes truthtelling and good faith [7]. Those working in practical ethics tend to argue that a professional should never lie, although complete candidness is not advisable either (for example, in medical treatment) [8]. Online mendacity has even been described in pathological terms. A syndrome called 'Munchausen by Internet' "occurs when medically well individuals fake recognized illnesses in virtual environments, such as online support groups" [9].

This is also a position enthusiastically adopted by social media sites, although whether they are driven by ethical considerations or the need to harvest quality data for their surveillance-based business models is perhaps moot. A brief scan through terms and conditions of major sites finds a touching devotion to truth and honesty. Facebook insists that "Facebook users provide their real names and information, and we need your help to keep it this way." So, for example, users "will not provide any false personal information on Facebook" and "will not create more than one personal account," and hence a Facebook account becomes identifying. Furthermore, "you will keep your contact information accurate and up-to-date."[1] Beyond your volunteering information, it will bombard you with questions to which it expects true answers. Twitter insists you don't impersonate others in a manner that misleads or confuses others (and makes it clear that your intention to mislead, though a confounding factor, is not necessary).[2] Misleading is a cardinal sin on Twitter, and many prohibitions are characterized by its presence. This rather belies its promotional claim that "you are what you tweet!"[3] It is perhaps more that "you must tweet what you are!"

2.1 The Practice of Online Mendacity

Only extreme individualists such as Nietzsche [10] and Max Stirner [11] are prepared to make the case that norms of truthtelling are harmful. Stirner writes of the "heroism of the lie." Yet the fact that hi-tech fibbing is frowned upon does not mean that people do not do it. There is a wide literature here [12], and some of the current authors conducted an empirical investigation into ordinary practices of fabrication and omission not for personal financial gain, but as a technique in life management and informational self-determination [13].

Over half of the respondents to the survey admitted to lying at least sometimes, though relatively few admitted to doing it often. The survey also attempted to elicit reasons and strategies for lying, which broadly revealed desires to retain some influence over their online experience, based on a perception that subjects had sacrificed control data-hungry Internet Behemoths.

Reasons for economy with the truth included:

- *playing up* events to make them funnier or more impressive and *playing down* problems: "Lied about my mental health countless times, denied depression and suicidal thoughts."
- *Privacy* and *mistrust* of the system were also motivations, to stop online identity being connected to a 'real' identity, and the collapse of contexts between different online services: "More than hiding my identity, it is a way … to prevent such platforms from connecting together my different identities, and then jumping to conclusions I did not ask them to make."
- People also managed their identity to *conform* or avoid discrimination: "The major untruth I tell is pretending to be a man rather than a woman on YouTube – I know it's bad and not helping the cause, but I know that if I want to convince someone of a particular point, if I pretend to be a man my sayings won't be regarded through the bias of my gender."
- The fluidity of online identity was also used for *roleplay* and *experiment*: "I have created two alter-egos. One was a short-lived novelty account that posted in the voice of a fictional character, while the other is a member of a hate group whom I used as a kind of psychological experiment in empathy--by performing as a member of that group, I came to a fuller understanding of what compels their bigotry."

None of these strategies is essentially ignoble or crooked. Many of the motivations are responses to power imbalances, either between social groups or with service providers. Others are known strategies for managing relationships with others, adapted for contexts mediated through social networking sites which themselves are not neutral players in the game.

2.2 Privacy With Porky Pies

Amending data to protect privacy has been accepted for a long time. Anonymization of data includes aggregation, and removal of direct identifiers, but perturbation techniques (e.g. Barnardisation) are also used to falsify the record while leaving desired statistics (e.g. the mean or variance of a population) unaffected.

Noting that self-surveillance is increasing, Dodge and Kitchin argue that it is important for the individual to inject uncertainty into the capture, storage and transfer of data when she has control [14]. They propose not only deletion and degrading precision, but actual falsification, including specific and random misrecording of (some of) the details of an event, rescripting certain events after a period of time, and biasing the data from an event in line with a generic standard. The point of doing this is to protect the data subject's privacy; even if the data was used by someone else, they couldn't be as confident that they correctly represented the actions or choices of the subject.

Dodge and Kitchin argue that not only is this not wrong, it is an ethical way to go about privacy protection. Data gathering is not a neutral process producing an Archimedean view of the world; interests and incentives are built into it. Given that, there is a case for embedding the interests of the data subject into the process.

3. SELF-SURVEILLANCE AND PERSONAL DATA MANAGEMENT ARCHITECTURES

In line with Dodge and Kitchin's expectations [14], we are increasingly indulging in self-surveillance and quantifying the self, gathering and gamifying health data with wearable devices, storing to-do lists, using activity trackers, repatriating personal data via government programmes such as Midata [15], and

[1] https://www.facebook.com/legal/terms, in clause 4.

[2] https://support.twitter.com/articles/18311-the-twitter-rules#.

[3] https://twitter.com/tos.

leaving rich traces on social networks [16]. Privacy concerns are exacerbated by the increased tendency to share information with our networks, [17] a trend not anticipated by Dodge and Kitchin

But this tendency brings risk. As technology improves, data that it is safe to share now may not be safe in the future. Templeton uses the analogy of 'Time travelling robots from the future': the information collected now will be subjected to increasingly sophisticated analysis techniques over time. For example, it may become possible to carry out face recognition on CCTV footage to reconstruct the movements of a large number of citizens [18]. Data decays unpredictably, so one cannot know what will be available and what deleted in 10, 20 or 50 years' time. The context, including law and social norms, will evolve. Data will be bought and sold hundreds of times over. In a situation of such dynamic uncertainty, how can we restore a semblance of control?

3.1 Re-De-Centralizing the Web

The Web was designed as a decentralized information and communication tool; anyone could link to, or download, anything, without referring to a central gatekeeper. Recently, this model has been frayed by major players who, via the economic forces of network effects, technological lock-in and low marginal costs of adoption, can amass giant user bases for their walled gardens. As data is not portable, a move to an alternative platform is unattractive. One would have to build up one's data from scratch, on a smaller and therefore less valuable network [19].

Web users now resemble babies with a tenuous hold on their personal data candy. The current model of exploiting big data is alienating. Data is harvested from users and consolidated in giant databases where analytics produce monetizable insight to the benefit of data gatherer-owners, while the data subject gets a better class of spam. People are decoupled from their data, unable to manage, curate or police it, facilitating abuse, including use for unintended or unconsented-to practices and irresponsible handling and storage. Identity partitioning is hard, and often ruled out by the terms and conditions of the major walled garden companies. Identity consolidation is preferred, positioning the major platforms as central information controllers able to link across contexts to generate a rich picture of a person's activities. The Snowden revelations provide a dark context for this unaccountable big data paradigm, and it is unsurprising that trust in big data is at a relatively low ebb [20].

One class of technologies with the potential to rebalance some asymmetries and restore some trust are architectures which allow the data subject some measure of control over, or input into the exploitation of, her personal data, including the data she has collected herself (via lifelogging, for example), and data collected or inferred about her. We call these *Personal Data Management Architectures* (PDMAs), and intend this term to be agnostic over particular architectures, affordances and business models. It includes, but is not restricted to, Personal Data Stores (PDSs) and Personal Information Management Services (PIMS) [21], [22], [23], [24]. The services PDMAs might provide include user-centric consent management tools, preventing external access to data except under approved conditions, negotiating privacy policies, or even allowing access to rich sources of data from self-surveillance for payment, free services or other benefits. It is important to note that such services *do not* depend on the PDMA storing data, and we make no assumption that they will necessarily provide storage services (although PDSs do).

The PDMA could act as a privacy and identity assistant, with an understanding of context (such as interaction history), mapping multiple identities to different activities, and establishing trust credentials from those requested access to the data. Forced identity consolidation would no longer be appropriate, and data would have some portability across at least some contexts. The PDMA would manage interactions so that external parties would not be aware that, say, the employee of a well-known bank, the player of World of Warcraft, the denizen of a fetish site and the campaigner for immigration rights are all the same person. The analytics would still enable societal benefit from big data, but the benefits would be more widely shared. In-house number crunching would be replaced by distributed querying of distributed data stores.

PDMA technology is not mature, but in this position paper we assume that innovation capable of providing the above-mentioned services could happen in the near future. Assuming a mature market of critical mass emerges, PDMAs would subvert the current big data regime. If data subjects had greater powers to control the use of data about them, it would be far simpler to attach provenance to it, and hold people accountable for its use and the decisions made on the back of it. If they were concerned about their privacy, autonomy, or merely sharing the benefits from the use of their data (which currently accrue entirely to others), data subjects would have a route to engagement, and incentives to share richer and more personal data with companies for defined mutual benefits.

In effect, the Web, which is centralizing around the major platforms, would be re-de-centralized by socially-aware PDMAs. For the purposes of this paper, we assume this happy outcome, but for a defense of its feasibility see [24].

3.2 21st Century Devious Man

PDMAs can be used to provide (some) misleading information to protect privacy; for details see [25]. Several strategies are possible. As data is given out, it could be perturbed systematically to create inaccurate but privacy-preserving effects.

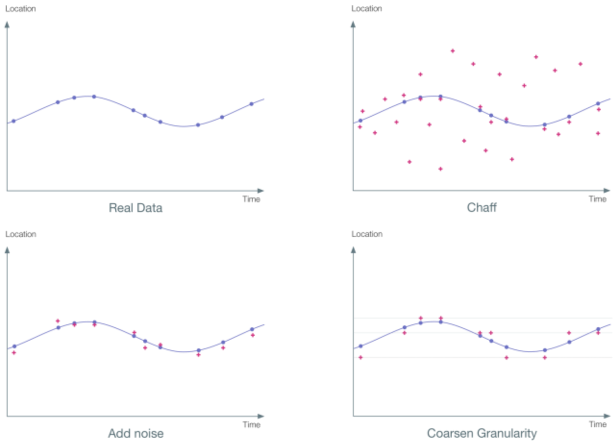

Figure 1: Three Obfuscatory Strategies

For example, a collection of accurate data could be perturbed using a number of methods. Extra spurious data (*chaff*) could be added to the correct data. Some noise could be added, as with Barnardisation. The granularity of the data could be coarsened (see Figure 1 for graphical representations of these three). Some deviations could be made systematically – for instance, change data implying I was at the pub to data implying I was in the library. The system might infer a normal pattern of behaviour for, say, Wednesday afternoons, and replace an abnormal Wednesday with data implying a generic one (or, alternatively, it could do the opposite and replace data from a normal day with something more unusual). Or, one could combine these to create multiple overlapping traces, each locally coherent and plausible, but not simultaneously consistent (elsewhere we have called this strategy *palimpsestification*). Each of these strategies is possible, many are actually available on some applications, and some are standard anonymisation practice [25].

The techniques available will be enhanced if data is shared reciprocally with friends so that collusion can aid in the creation and verification of otherwise unachievable data streams. For example, one could present data from a friend as one's own, thereby being apparently co-located with them. One's friends could add obfuscatory data to their own PDMAs that confirms one's story. Or several people could collaborate on an account, making it hard to identify each individual contribution (see Figure 2 for graphical representations of these). These ideas are more speculative, as few computational systems of this type exist. There are aspects which make these strategies harder to pull off, as coherence is required across multiple different accounts, but if successful, the obfuscation would be better supported and harder to detect [25]. The social aspect to the obfuscation is consistent with the direction of travel in user-centric data management [17].

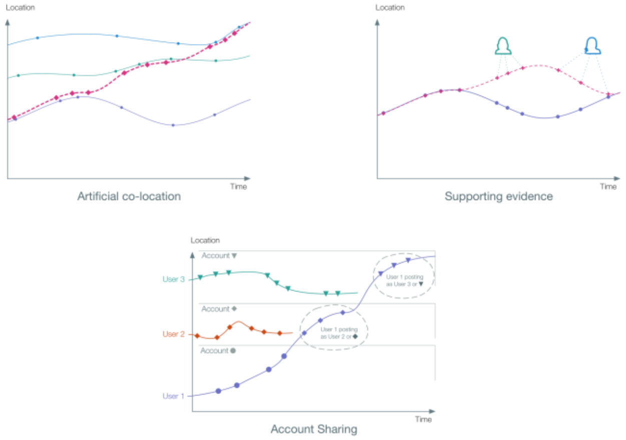

Figure 2: Three Collaborative Obfuscatory Strategies

There are of course major ethical issues implied by such practices. These are not discussed here, but we do not underestimate them; for a full discussion, see [25]. However, it is worth noting that under current data protection legislation, the use of personal data requires it to be depersonalized, often using anonymisation techniques analogous to the ones discussed here. Hence this is not unexplored territory.

4. IT'S ALL IN THE CONTRACT

In this section, we will heroically assumpe that the PDMA ecosystem sketched in section 3 is functional, and consider some legal implications, focusing specifically on UK law.

4.1 PDMAs and Contract

A PDMA would mediate data interactions between a user and a data consumer. Current privacy policies act as contracts between data subjects and data gatherers, and clicking 'yes' on the privacy button establishes a contract between gatherer and subject (leading to the use of the disparaging term 'clickwrap').[4] Although the US Federal Trade Commission has built up a rich legacy of case law (e.g. [26], [27]), the position is hardly ideal, given that privacy policies are usually unread and (deliberately?) complex, and that situations can change after acceptance, particularly with respect to the content of the privacy policy, without requiring a new contract.

It would make sense, therefore, to expect the mediated interaction between PDMA and data consumer to be governed by a contract as well, except that in this case the user (via the PDMA – one of its functions) would negotiate with the consumer as to its form. If the demand for the data was high, the user might be able to make a take-it-or-leave-it offer, rather as data gatherers do now. If not, then the negotiation would be more complex, and given the evolution of a market for services and tools, we might expect such negotiations to become well-rehearsed over time. Ultimately, the legal question will rest on what is in the contract and how it is drawn – it's all in the contract.[5]

That is not to say that there will not be many complexities as the PDMA makes an offer of terms, and the data consumer contests them. Will there be a battle of forms, with each side coming up with its own terms, and how would a negotiation take place? There would need to be automated mechanisms that would both scale and produce immediate results. In that event, assuming little or no human involvement in the negotiation process beyond setting targets, policies and red lines, there would be a question as to the validity of the contract produced. Obvious issues would include rights over derived data, or aggregated data, use of third party services, and which country's laws should regulate the storage and processing of the data.

4.2 Data Protection, Privacy and Obscurity

The secrecy paradigm of privacy [28], based on concealment so that disclosed information can no longer be private, fails to recognise that individuals want to keep things private from some but not others and that context is a central factor [6]. The means with which a user employs privacy protection will often determine whether it is public or private – mechanisms such as passwords or encryptions suggest an attempt to keep information private and thus eligible for protection. There is not yet consistency in courts regarding the extent to which context will be considered relevant.

PDMAs require a new set of arrangements between data subjects and consumers, with repercussions about how data and privacy are understood. In particular, there is currently a debate about the effects and desirability of the commodification of personal data; in particular, would commodification lead to irresistible threats to privacy? Schwartz has argued that for a market for personal data to be properly privacy sensitive, it requires *inter alia* restrictions on what data can be transferred (inalienabilities). Beyond natural caveats, such as access for law enforcement, Schwartz suggests that "the ideal alienability restriction on personal data … would permit the *transfer* for an initial category of *use* of personal data, but only if the customer is granted an opportunity to block further transfer or use by unaffiliated entities" [29]. Furthermore, the 2009 judgment of the European Court in Reklos v Greece

[4] http://www.clickwrapped.com/.

[5] Thanks to David Chalk for this pithy statement of the principle.

(application no.1234/05) grants the privacy rights of Article 8 of the European Convention on Human Rights in the context of someone wishing to retain control of his image. It is not a stretch to imagine this implying similar rights over personal data.

Other commentators argue that such arrangements pose a threat to the public domain. Information is in the public domain except when it is covered by a specific carveout such as copyright, and so facts about an individual should generally be placed in the public domain. Undoing this assumption may reduce societal benefits. Conversely, where information is kept from the public, for example when gathered by an e-commerce firm, this often provides economic benefits. Undoing these arrangements could severely damage markets, by undermining the surveillance/free service/advertising model [30].

These are factors to take into account, and a reason to move in baby steps. However, the PDMA system suggested here need not have this catastrophic effect. Firstly, though firms would encounter new obstacles, they may also gain through access to other, richer sources of information (for example via quantified self applications that the subject is willing to share in return for services). Secondly, the negotiations between a subject's PDMA and the data consumer would allow a defence of the consumer's monetization model against the PDMA's privacy policies.

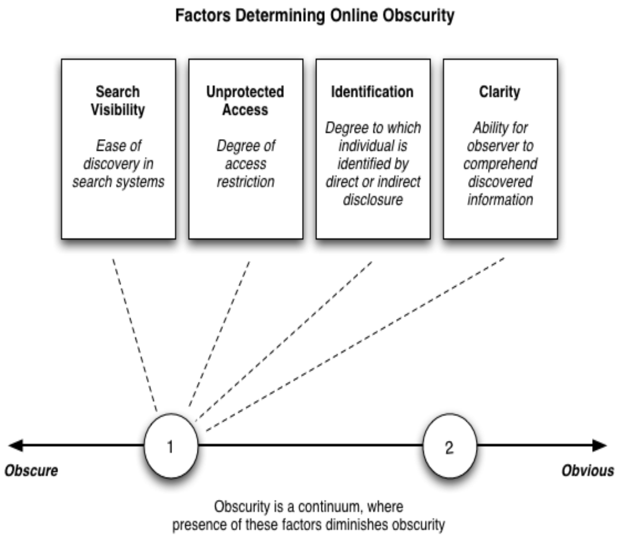
Factors Determining Online Obscurity
Search Visibility
Ease of discovery in search systems
Unprotected Access
Degree of access restriction
Identification
Degree to which individual is identified by direct or indirect disclosure
Clarity
Ability for observer to comprehend discovered information
1
2
Obscure
Obvious
Obscurity is a continuum, where presence of these factors diminishes obscurity

Figure 3: Four Factors Determining Obscurity [31]

Another means of providing a more context-sensitive view of privacy in the courts is to follow up the suggestion of Hartzog and Stutzman, based on ideas of contextual integrity [6], that information that is less sensitive or sensitive in fewer contexts should be protected by obscurity if it is ineligible for more robust privacy protections. On their proposal for law reform the courts would take into account a continuum of obscurity, "a framework whereby protection with varying strengths exist and considered cumulatively fall along a spectrum that will allow the courts to make a more nuanced analysis of online information on a scale of obscurity" [31]. Information is obscure online if it exists in a context missing key factors essential to discovery or comprehension, such as search visibility, unprotected access, identification, and clarity. The protections provided by a PDMA would affect the position of information along the continuum, and reduce onerous constraints on data consumers where the PDMA was most effective.

4.3 Could Obfuscation Be a Criminal Offence?

Finally, it is important to consider that in many circumstances, obfuscation might be a criminal action. There is concern in law enforcement and security circles that privacy protection could provide cover for many illicit projects from paedophilia and terrorism down to fraud and trolling. The automated nature of many PDMA exchanges would be no defence, as, for example, the Fraud Act 2006 covers instructions given to an automated machine. In addition, the content of terms and conditions may link to the risk of a computer misuse offence being committed under the UK's Computer Misuse Act 1990.

Common factors in judgments in this area are the conduct of the data subject and the value of the information that would be obscured. The conduct and the motivations of the subject are typically a decisive factor as to whether they were expressing rights of freedom of speech or association. Conversely, a user being defamatory or dishonest, or trying to cause a loss for the data consumer, is likely to receive harsher legal repercussions.

Obfuscation may be both `pro-social' and antisocial. Systematic untruth can weaken the social fabric, disrupting trust with friends and colleagues, as well as those carrying out legitimate surveillance. This adds friction to interaction, requiring increased resources for verifying data. More broadly, the social good which comes of having access to increasingly detailed personal data can be compromised if significant proportions of the data are untrue [25]. Obfuscation is a positive act of choice and concealment, and as such must be located in a framework of accountability.

5. CONCLUSION

To conclude, we believe that there is sufficient traction in the model of PDMA-mediated data curation and strategies for obfuscation to merit further study. That is not to underestimate the technical, legal and jurisdictional challenges. But the autonomy and privacy of data subjects would be enhanced, which will ultimately have a positive effect on trust in online transactions and e-commerce. Whether this positive effect will fully offset the losses from the withdrawal of the current anything-goes model of big data is a moot point. It is also worth pointing out that regulation in this area may let society off the hook of evolving norms for constraining surveillance-based business models to the satisfaction and profit of all.

6. ACKNOWLEDGMENTS

DMR, KOH, MVK and NS are supported by SOCIAM: The Theory and Practice of Social Machines, funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/J017728/1. The authors are grateful to Chloe Murthwaite and Bradley Meaker of the University of Winchester for their assistance in the research for this paper.

7. REFERENCES

[1] Government Office for Science, 2013. *Future Identities: Changing Identities in the UK: the Next 10 Years*, Foresight report, https://www.gov.uk/government/uploads/system/uploads/atta

chment_data/file/273966/13-523-future-identities-changing-identities-report.pdf.

[2] boyd, d. 2002. *Faceted Id/Entity*, MSc thesis, Massachusetts Institute of Technology, http://smg.media.mit.edu/people/danah/thesis/danahThesis.pdf.

[3] Lyon, D. 2001. *Surveillance Society: Monitoring Everyday Life*, Open University Press, Buckingham.

[4] Butler, J. 1997. *Excitable Speech: A Politics of the Performative*, Routledge, New York.

[5] Van House, N.A. 2011. Feminist HCI meets Facebook: performativity and social networking sites. *Interacting With Computers*, 23, 5, 422-429.

[6] Nissenbaum, H. 2009. *Privacy in Context: Technology, Policy and the Integrity of Social Life*, Stanford University Press, Palo Alto.

[7] Habermas, J. 1984. *The Theory of Communicative Action Vol.1: Reason and the Rationalization of Society*. Polity Press, Cambridge.

[8] Jackson, J. 2001. *Truth, Trust and Medicine*. Routledge, London.

[9] Pulman, A., and Taylor, J. 2012. Munchausen by Internet: current research and future directions. *Journal of Medical Internet Research*. 14, 4.

[10] Nietzsche, F. 1996. *Human, All Too Human: A Book for Free Spirits*. Cambridge University Press, Cambridge.

[11] Stirner, M. 1993. *The Ego and its Own: The Case of the Individual Against Authority*. Rebel Press, London.

[12] Hodder, M., Churchill, E. and Cobb, J. 2013. *Lying and Hiding in the Name of Privacy*, Customer Commons Research, http://customercommons.org/wp-content/uploads/2013/05/CCResearchSurvey1Paper_Final.pdf.

[13] Van Kleek, M, Murray-Rust, D, Guy, A., Smith, D.A. and Shadbolt, N.R. 2015. Self-curation, social partitioning, escaping from prejudice and harassment: the many dimensions of lying online. In *Proceedings of the 2015 Web Science Conference (WebSci15)*, Oxford, UK.

[14] Dodge, M. and Kitchin, R. 2007. 'Outlines of a world coming into existence': pervasive computing and the ethics of forgetting. *Environment and Planning B: Planning and Design*, 34, 431-445.

[15] Shadbolt, N. 2013. Midata: towards a personal information revolution. In Hildebrandt, M., O'Hara, K. and Waidner, M. (eds.), *Digital Enlightenment Yearbook 2013: The Value of Personal Data*. IOS Press, Amsterdam, 202-224.

[16] Whitsun, J. 2013. Gaming the quantified self. *Surveillance and Society*, 11, 1, 163-176.

[17] O'Hara, K., Tuffield, M.M. and Shadbolt, N. 2009. Lifelogging: privacy and empowerment with memories for life. *Identity in the Information Society*, 1, 2, 155-172.

[18] Templeton, B. n.d. *A Watched Populace Never Boils*, blog, http://www.templetons.com/brad/watched.html.

[19] Zittrain, J. 2008. *The Future of the Internet: And How to Stop It*. Yale University Press, New Haven.

[20] Coll, L. 2015. *Personal Data Empowerment: Time for a Fairer Data Deal?* Citizens' Advice Bureau, London, https://www.citizensadvice.org.uk/Global/CitizensAdvice/Corporate%20content/Publications/Personal%20data%20empowerment%20report.pdf.

[21] Nguyen, M-H.C., Haynes, P., MacGuire, S. and Friedberg, J. 2013. A user-centred approach to the data dilemma: context, architecture, and policy. In Hildebrandt, M., O'Hara, K. and Waidner, M. (eds.), *Digital Enlightenment Yearbook 2013: The Value of Personal Data*. IOS Press, Amsterdam, 227-242.

[22] Heath, W., Alexander, D. and Booth, P. 2013. Digital Enlightenment, Mydex, and restoring control over personal data to the individual. In Hildebrandt, M., O'Hara, K. and Waidner, M. (eds.), *Digital Enlightenment Yearbook 2013: The Value of Personal Data*. IOS Press, Amsterdam, 253-269.

[23] Ctrl-Shift 2014. *Personal Information Management Services: An Analysis of an Emerging Market*, Nesta, London, http://www.nesta.org.uk/sites/default/files/personal_information_management_services.pdf.

[24] Van Kleek, M. and O'Hara, K. 2014. The future of social is personal: the potential of the personal data store. In Miorandi, D., Maltese, V., Rovatsos, M., Nijholt, A. and Stewart, J. (eds.), *Social Collective Intelligence: Combining the Powers of Humans and Machines to Build a Smarter Society*, Springer, Cham, 125-158.

[25] Murray-Rust, D., Van Kleek, M, Dragan, L. and Shadbolt, N. 2014. Social palimpsests: clouding the lens of the personal panopticon. In O'Hara, K., Nguyen, M-H., and Haynes, P. (eds.), *Digital Enlightenment Yearbook 2014: Social Networks and Social Machines, Surveillance and Empowerment*. IOS Press, Amsterdam, 75-98.

[26] Federal Trade Commission 2013. *FTC Approves Final Order Settling Charges Against Epic Marketplace, Inc*, press release, https://www.ftc.gov/news-events/press-releases/2013/03/ftc-approves-final-order-settling-charges-against-epic.

[27] Federal Trade Commission 2014. *FTC Approves Final Order Settling Charges Against Snapchat*, press release, https://www.ftc.gov/news-events/press-releases/2014/12/ftc-approves-final-order-settling-charges-against-snapchat.

[28] Solove, D.J. 2006. A taxonomy of privacy. *University of Pennsylvania Law Review*, 154, 3, 477-560.

[29] Schwartz, P.M. 2004. Property, privacy and personal data. *Harvard Law Review*, 117, 7, 2055-2128.

[30] Lemley, M.A. 2000. Private property. *Stanford Law Review*, 52, 1545-1557.

[31] Hartzog, W and Stutzman, F. 2013. The case for online obscurity. *California Law Review*, 101, 1-49.