On the Ethics of Using Publicly-Available Data

Antony K. Cooper, Serena Coetzee

On the Ethics of Using Publicly-Available Data

Antony K Cooper[1,2][0000-0001-9411-2094], and Serena Coetzee[2][0000-0001-8683-8047]

[1] Smart Places, CSIR, PO Box 395, Pretoria, 0001, South Africa
[2] Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa
acooper@csir.co.za, serena.coetzee@up.ac.za

Abstract. Publicly-available mobile data can be used to derive fine grain commuting and travel patterns. These types of data include geocoded or geotagged discrete units of communication: messages, posts, tweets, status updates, check-ins, images and the like on a variety of social networking services. Clearly, there are ethical issues concerning the use of such data, particularly the invasion of privacy. A review of the literature has been done to explore these issues.

Keywords: Privacy, Surveillance, Mobile data, Social Networking Service.

1 Introduction

The more one knows about people and their condition(s), actions, needs, preferences, beliefs and the like, the better one can provide services to them – if one is in a purely benevolent, altruistic, caring and diligent society. However, there are problems or ethical issues with such data, and with capturing, the processes for capturing, and using the data. Even then, interpretation of the data might be invalid.

As a concept, *publicly-available data* usually refers to data found readily (such as on the Internet) and accessed (downloaded) easily and for free. Many of these data sets are created and distributed by public organisations. Publicly-available data includes open data (freely usable, reusable and redistributable without restrictions), data available on request, public-domain data (without copyright), copyrighted data, commercially-sold data and data with limited availability (eg: for a limited time or for only specified uses). Clearly, any data could fall into more than one of these types.

This paper draws on PhD research and a project that investigated using publicly-available mobile data to derive commuting and travel patterns. Clearly, there are ethical issues over using such data, particularly invading privacy. We review and analyse the literature to explore the ethics of using publicly-available data. We do not attempt to pick a framework of normative ethics on using such data. Rather, we explores some of the issues, focusing on surveillance-type data and hence on privacy.

1.1 Characteristics of publicly available data

The following are some characteristics of publicly-available data.

- **Surrogates**: digital data are not the real world, but merely represent the real world, and often because of costs, availability, laws and so on, the data actually recorded are merely a surrogate (or proxy or approximation) for the phenomena in the real world that are meant to be measured or assessed.
- **Big data**: it is easy to capture vast quantities of data, but often difficult to extract meaning or to forewarn from the overwhelming data. Even worse, some assume the sheer volume provides greater objectivity, neutrality and accuracy, but "*data are always the result of conscious, subjective decisions on the part of researchers, and are the result of inherently social processes*" [1]. The perceived authority or effectiveness is lost by being overwhelmed by all the automatically-collected data, by mistaking omniscience for omnipotence or intelligence: "*the more you know about the secret lives of others, the less powerful you turn out to be!*" [2]
- **False precision**: digital geographical coordinates are often given to some arbitrary precision based on a data storage decision (eg: single vs double precision), rather than the accuracy of the recording method, giving incorrect perceptions of coordinate accuracy. For example, a point geocoded from a toponym (eg: Tshwane) could have a precision of a second (about 30m), though the toponym encompasses many square kilometres. False precision also applies to other types of data.
- **Quality**: many factors can inhibit the quality of data, but these are often not well understood by the end users.
- **Metadata**: documenting data, their quality and characteristics are essential for being able to use the data meaningfully, but metadata is often not well understood by the end users and is often not provided adequately.

1.2 Potential ethical issues with publicly-available data

Whatever the nature of the data, the following are some issues (which have ethical aspects) with the creation, distribution and use of publicly-available data.

- **Privacy**: there are moral and legal concerns over the invasion of the privacy (or **surveillance**) of individuals, groups and organisations, and these are the main focus of this paper.
- **Bias**: because of the above and subjectivity in deciding what attributes to collect and how, any data set is invariably a biased representation of the population. While this can be ameliorated through other data, local knowledge and insights, and careful statistical analysis, it is of particular concern when those using the data are blissfully unaware of the bias. Bias also occurs in training sets for models, such as in machine learning. Error is ubiquitous [3].
- **Liability**: this could be for incorrect data, which then compromises someone's rights, endangers safety and security, wastes money and other resources, and so on.
- **Right to exploit content**: on the other hand, for some (such as entertainers and artists) it is important to be able to exploit their content publicly which might be inhibited by corporations controlling the content – analogous to censorship.
- **Censorship**: this can be disguised and rationalized as prudent selection, due to the limited budget of a public library, to suppress hate speech, to maintaining literary excellence, to ensure balance and/or to meet the audience's requirements [4, 5].

Invading privacy (or surveillance), censorship and liability are often used as excuses for one another. For example, content could be denied or restricted (censored) to "protect" privacy or because of "concern" over liability. Further, claims over content ownership are used to censor content or restrict use, frustrating creators: Toya Delazy released her album online as her record label was limiting stock availability [6].

On the other hand, privacy could be compromised over "concerns" over liability, such as when a company monitors staff emails. The issues are not well understood either, such as when "poor" data (eg: low resolution remotely-sensed imagery) is considered to be censored data, because it covers in inadequate detail, an area of interest to a conspiracy theorist, or the like. However, "*privacy and security do not have to contradict each other; indeed, secure online interactions, enabled by a secure online identity, is a precondition for full internet freedom*" [7].

These issues also apply to private or restricted data, as inappropriate surveillance or data exploitation can be done within limited or closed groups.

2 Ethics

Ethics concerns the nature of ultimate value and the standards by which human actions can be judged right or wrong. Ethical judgement is influenced by the values of a person or group: their convictions of what is good or desirable [8]. Values are determined by different factors, including culture, religion, social and economic status, personal experiences, age, gender and profession. Data are often shared globally and the ethics of using such data is subject to significantly diverse value systems. **Normative ethics** aims at establishing the norms or standards for appropriate conduct and applied ethics is how these are used to deal with practical moral problems. There are three major approaches in normative ethics, which in practice, are often mixed:

1. **Virtue ethics** emphasises virtues or moral character as a way of assessing or justifying each and every action or non-action.
2. **Consequentialism** emphasises the consequences of actions, which can be interpreted as the end justifies the means; and
3. **Deontological ethics** emphasises duties or rules, which can be reduced to a check list of what to do in different situations [8, 9].

Deontological ethics is perhaps the easiest to adhere to in practice, because in each situation, one can look up what is the appropriate thing to do. Essentially, legislation is a form of deontological ethics. However, problems with deontological ethics are:

- Someone can use them without having any moral understanding of exactly what they are doing (or not doing) and the implications thereof;
- If there is no obviously applicable rule in a particular situation, the person has no meta-framework or set of values to use to decide on the best course of action;
- Reciprocity can be difficult as the values or rules one person uses for determining how to behave towards another might be incompatible with those used by the second person, causing conflicting understanding of the actions and reactions; and

- Without a meta-framework, people will tend towards the softest option and/or try to push the boundaries of what is acceptable [8, 9].

Virtue ethics focuses on the moral character and the need to educate and develop such a moral character. Considering what a 'virtuous person' would do can guide ethical decision-making [8]. There are various forms of consequentialism, such as utilitarianism (good conduct has consequences that achieve the greatest good to the greatest number of people) and situational ethics (considers the context in which conduct takes place and the consequences within this context).

Artificial intelligence and other sophisticated tools can be used to identify ethical and unethical behaviour, such as on social media, and assess the veracity of news stories and images [10]. Such tools can also be used unethically and to create and disseminate fake news. One needs to consider how these tools function and their outputs, to embed robust ethical analysis and decision making in the tools [11].

When conducting research that collects private data, one obtains *informed consent* to invading someone's privacy and publishing the research results, as part of a research ethics process. What constitutes informed consent is in itself an interesting problem in ethics, due to language, literacy, education, coercion, rewards, etc.

A problem with informed consent is that the research subject needs to remember what they have agreed to and when. Unfortunately, this is not always the case, as we found in a project tracking participants to and from an event [9, 12]. If the user has to opt-in to the tracking, there is likely to be a high loss rate. If the user has to opt-out, they might forget to stop the tracking. Such issues of informed consent apply to private data obtained for government, commercial and other purposes, which are often obtained without a formal ethical review and might have the *informed* part buried in fine print and the *consent* part implicit rather than explicit.

3 Privacy and protecting privacy

"*The right to life has come to mean the right to enjoy life, – the right to be let alone; the right to liberty secures the exercise of extensive civil privileges*" [13]. Their primary concern was over making private details public: "*each crop of unseemly gossip, thus harvested, becomes the seed of more, and, in direct proportion to its circulation, results in the lowering of social standards and of morality*" [13]. Further, "*it is also immaterial that the intrusion was in aid of law enforcement*" [14].

Perhaps the antithesis of data democratization and freedom of information is making too much available, compromising the privacy of individuals especially, but also of organisations. Privacy is complex to define, being perceived differently by different cultures and treated differently in legislation. Privacy is perceived as being about protecting people's personal information, but it also includes territorial (or location) privacy, physical (or bodily or health) privacy and privacy of communications. Privacy is not the same as confidentiality or secrecy, though they can overlap [15].

Many sacrifice their privacy voluntarily, especially when using social media, but they could be doing so through ignorance, deception, coercion or peer-pressure. Unfortunately, social media sites are notorious for changing privacy settings (sometimes

through “errors”) and/or for making them complex. Even when personal data are secured in a private area, they could still be exposed through changes in legislation, decisions by courts (eg: search warrants) and company buy-outs.

Many governments have introduced legislation to protect privacy to varying extents. Perhaps the best known and most significant because of its wide applicability is the European General Data Protection Regulation (GDPR), which came into effect on 25 May 2018 [16]. The principles of the GDPR are lawfulness, fairness and transparency; purpose limitation; data minimisation; accuracy; storage limitation; integrity and confidentiality; and accountability. The South African equivalent to the GDPR is the Protection of Personal Information Act (POPI) [17].

4 Invasion of privacy

“*Privacy is mostly an illusion. A useful illusion, no question about it, one that allows us to live without being paralyzed by self-consciousness. The illusion of privacy gives us room to be fully human, sharing intimacies and risking mistake*” [18].

4.1 Covert surveillance

Covert surveillance is possibly what many consider surveillance to be: monitoring behaviour and communications surreptitiously, for detecting, investigating and monitoring threats (criminal, terrorist, social unrest, etc), influencing and controlling society, and, hopefully, protecting citizens. For example, “brain fingerprinting” is claimed to detect the presence or absence of information in someone’s brain, using electroencephalography (EEG) [19], though there are concerns over the studies [20].

4.2 Trans-jurisdiction surveillance

One feature of the designed-in robustness of a packet-switching network such as the Internet, is one cannot guarantee the routing of individual data packets. Even with a high-speed, high-bandwidth Internet connection directly between two countries, parts of the connection might be routed through other countries – which might capture and/or study the data traffic *en route* [21]. Such trans-jurisdiction surveillance might be accidental; though those doing the surveillance should realise it happens. For example, Internet traffic to and from the United Nations in New York is presumably routed through the USA and hence likely to be recorded by the NSA. It appears that Internet traffic can be misdirected deliberately and surreptitiously, particularly across national boundaries, to inspect and/or modify the transmitted data [22].

Another example concerns virtual private networks (VPNs). They are used to ensure that anyone intercepting the (often encrypted) traffic cannot read what is being transmitted (or perhaps even where the source and destination are), but the traffic gets routed through servers, which lends itself to surveillance by the server owners.

Another form of trans-jurisdiction surveillance is remote sensing, with an early use of LANDSAT satellites being to monitor wheat crops in the Soviet Union [23].

4.3 Overt surveillance

Not all surveillance is covert, with overt forms including those visible and well identified (such as CCTV surveillance cameras in public, or disclaimers of a call being recorded) or to which one agrees explicitly (such as the small print for using a web site). However, in some jurisdictions, such supposed agreements might be unenforceable, being excessively long or changed arbitrarily and without notice [24].

Further, it is easy to forget one's actions are being observed, even when giving explicit consent [9, 12]. Clearly, this leads to complacency and the risk of becoming accustomed to the surveillance society. It is also easier to accept surveillance when under the influence of someone one trusts, such as parents recommending their children enable mobile phone location disclosure services [25].

4.4 Overloaded surveillance

Apparently, the American NSA "*intercepts and stores nearly two billion separate e-mails, phone calls, and other communications every day*", making the system too complex to determine if it actually works [26]. Rather than wisdom, the sheer volume creates information entropy – so information becomes noise as it "*is routinely distorted, buried in noise, or otherwise impossible to interpret*" [26].

Consequently, such agencies probably create their own filter bubbles, due to, not in spite of, the sheer volumes they harvest. Much of the content (facts, opinions, allegations, imagery, comments, conversations, etc) will be contradictory, so the selection, rating and analysis will be biased by preconceived notions and desire to "*simply want to believe something that feels right*" [26]. It is easy to be so enamoured with sophisticated and expensive technology the basics get forgotten, with tragic consequences, such as the Boston Marathon bombing [27] and Navy Yard shootings [28].

Being able to conduct surveillance over the Internet, or use it to interfere with the rights of others, or conduct information warfare over the Internet are all quite different from being able to control the Internet! The genie is out of the bottle and cannot be replaced. The Internet was designed to be robust (distributed, with data sent in small packets) and self-healing if any node broke [29]. As the Internet pioneer John Gilmore put it, "*the Net interprets censorship as damage and routes around it*" [30].

4.5 Becoming accustomed to the *surveillance society*

It is easy to forget one is being observed. This can result in acting carelessly whilst being observed and/or accepting the lack of privacy by becoming used to it, or even by expecting it. Americans have been accustomed to limits on their privacy for many years [18], realizing Bentham's idea of the Panopticon [31]. The *Panopticon* is a circular building with an *inspection house* in the middle from which a custodian could observe secretly the inmates (around the perimeter) who could not communicate with anyone. Foucault [32] invoked the Panopticon concept[1] as a metaphor for the tenden-

[1] Though Brunon-Ernst [33] suggests that Foucault distorted Bentham's philosophy.

cy of modern "disciplinary" societies to observe and attempt to "normalise" their citizens. "*The panopticon induces a sense of permanent visibility that ensures the functioning of power*" [34]. Unsurprisingly, this can lead to limited, or even curtailed, political and personal freedoms, and the loss of self-reliance [35]. Dobson & Fisher [36] took Foucault's metaphor further, identifying three "post-panoptic" models:

1. *Bentham's original concept*, which they consider to be the one Foucault used;
2. *Panopticism II*, in the form of the "Big Brother" type of surveillance of [37]; and
3. *Panopticism III*, technology tracking humans and their activities, such as cell-phone tracking [9, 12, 38], GNSS receivers, RFID[2] and geo-fences[3]. Crucially, the technology for Panopticism III is relatively cheap, effective and widely available to anyone, and not just well-resourced national security agencies.

The 1844 British postal espionage crisis concerned the Post Office opening letters at the behest of a foreign power. As the Law Magazine observed, "*the post-office must not only be CHEAP AND RAPID, but SECURE AND INVIOLABLE*" [39]. However, even though widely known and causing a 'paroxysm of national anger', it did not impact on the popularity of the Penny Post, which increased rapidly thereafter [39]. "*Snowden's revelations will have demonstrated that in practice, the web-surfing, texting and emailing public are indifferent to the risks they run to their privacy*" [39]. Similarly, Lanier [40] was concerned 2013 would be the year of *digital passivity*, when the cool gadgets (such as tablets running only applications approved by a central commercial authority) made us accept the commercial and government *surveillance economy*. Carr [35] fears privacy could be perceived as an outdated and unimportant concept inhibiting efficient transactions, such as socializing or shopping.

4.6 Mutual surveillance

The psychological and social effects of prevalent surveillance result in people being so intimidated by authority and/or so used to surveillance they conduct self-policing and can be forced or encouraged to spy on one another, extending easily, cheaply and significantly the surveillance reach of the authority, be it a government, the military, a corporation or any other type of organisation [32, 41].

4.7 Making data already in the public domain more visible

A common claim is that it is fine to put data online that are already in the public domain but otherwise difficult to access, such as documents and photographs in archives. However, that allows data matching. Such online content can also be accessed readily by anyone without revealing their interests, for example, using Google Street View to examine a neighbourhood, be it to find security weaknesses for targeting burglaries, stalking a resident, or mere curiosity. Similarly, much personal data are published, often unwittingly, in online genealogies.

---

[2] Radio frequency identification, small passive or active transponders.

[3] Virtual or conceptual geographical perimeter or barrier.

This could apply to archives themselves, though they have established procedures (file plans) for what can be archived, how, where, when, why and by whom. Archiving is complicated by legal issues such as copyright and technical issues such as accessing the deep Web, volatile communities, broken links and dynamic content [42].

Some assume naïvely that content made publicly available on the Web can be expunged permanently at a whim. The European Court of Justice decided that anyone has "*the right to be forgotten*" and can require search engines to remove pages from search results for specified terms [43], going against the advice of its own Advocate General [44][4]. This has obviously been used by the unscrupulous to hide their activities. Such pages are not deleted; they are just removed from searches.

As a result, legitimate reporting by respectable organisations such as the BBC gets proscribed, contravening the public interest [45, 46]. Essentially, this defames that article's author by declaring their work illegitimate. The search engine's operator has to decide what is a valid removal request, but that is inappropriate [47, 48].

Some applications use ephemeral data to (hopefully) protect privacy, that is, content deleted permanently after a specified time. Examples are SnapChat for photographs and Silent Circle for two-way transmissions of voice, email, video, etc. However, there is doubt that ephemerality can be enforced securely [49].

Web scraping or harvesting takes content from Web sites. Collecting can be targeted and pre-arranged, such as harvesting metadata and data from members of a collaborative system, for instance data providers in a spatial data infrastructure (SDI). Collecting can use well-behaved bots (as search engines do for indexing the Web), or simulated human access. This raises issues of copyright, such as the "Google Defense" case concerning thumbnails of images [50].

A search engine obviously does some form of Web scraping to locate the content first, before being able to provide the rapid search responses users expect. To return results as quickly as they do, search engines are not always accurate (particularly the results count) and there is much of the Web they cannot access [51].

4.8 Combining and processing available data

It requires much skill, intelligence and persistence to link together analogue data from diverse sources to find common threads, as good detectives do [52, 53]. Now, it is far easier to combine data from different sources using pattern recognition, artificial intelligence or other sophisticated tools (data matching, behavioural tracking, text analysing, data mining, linkage analysis, statistical analysis, spatial analysis and machine translation), exploiting fast hardware and huge and persistent digital data bases.

Most 'big data' analysis is not done to invade privacy, but to examine questions otherwise unexplorable, to understand human, physical and environmental behaviours in different contexts, and (hopefully) benefit society [54]. Unfortunately, an individual can be identified uniquely with very few data points, even coarse ones, such as

[4] The Court issues only one judgement and no dissenting opinions, and all deliberations Court are secret. As we pointed out to the Court's press office, this encourages bad law by forcing judges to support the majority opinion and protects incompetent judges from public scrutiny.

with cellular telephone use [55], power consumption of a mobile device [56] or renting public bicycles [57]. Personal traits can be gleaned from the digital footprints people leave on social media, which some exploit for trust and resilience modelling [58]. There are also services available for a fee to track a mobile telephone [59]. Hence, "*there is no such thing as anonymous online tracking*" [60].

4.9 Opting in vs opting out

To varying extents in different jurisdictions, one has limited control over how much of one's personal information is known, retained by others and/or shared. Sharing one's information (*opting in*) can provide access to services, opportunities or prizes[5], such as loyalty programmes (sharing personal and behavioural data for discounts or loyalty points), subscriptions to paid content, exposing one's resumé to potential (and hopefully desirable) employers, security services such as vehicle tracking, research collaboration or even friendships. Further, for some the *right of publicity* [61] is key for their profession and income, through exploiting their names, photographs, likenesses, recordings and the like – but only if they have consented and are remunerated appropriately. In many jurisdictions, one nominally can *opt out* of divulging one's private information, but even that explicit declaration gets ignored [62].

User-generated geographical data are known as *volunteered geographical information* (VGI). Some object to the term because data so collected might not be *volunteered*, but rather contributed, collected or harvested irrespective of whether the subject opted in, opted out, was even aware they were contributing their personal details, or had forgotten they were doing so. Harvey [62] suggests differentiating between volunteered (VGI) and contributed (CGI) geographical (or locational) information. Further, *truth in labelling* in the metadata following pragmatic ethics would explain the provenance of the information, allowing assessment of its *fitness for use* and if the quality of the data has been compromised by lax standards or even malfeasance [62].

4.10 Assuming one has nothing to hide

For anyone who lived through Apartheid (or communism, fascism, etc), it should be obvious that everyone has something to hide from a repressive government. Even in a reasonably open and stable democracy such as the USA, an innocent person has the right to remain silent [63], and keep their matters private. "*The skeptics no doubt have noticed that governments are made up of people and that people are prone to misuse information when driven by greed or curiosity or a will to power*" [18].

Examples of ripostes to those justifying surveillance are: show me your credit card details; show me yours first; none of your business; and those with nothing to hide don't have a life [64]. The person wanting to protect their privacy does not have to justify their position: the person wanting to invade someone's privacy needs to justify it first [64]. The metadata of one's communications can also reveal personality traits,

[5] Which is why there are so many competitions out there, because they are a cheap way to harvest personal data that are up to date.

religion, politics, habits, movements, condition, relationship issues, etc [65]. It is not only keeping 'facts' about oneself private, but also the assumptions made about us from the available data [66]. Further, there is the problem of identity theft.

4.11 Legal complexities

Human beings need space where they are guaranteed to be free from surveillance or interference by anyone, such as to establish and preserve intimate human relationships and develop intellectual faculties through reading, private conversation or writing privately [67]. It is very difficult to grow intellectually if one cannot experiment with ideas without fear of surveillance and resulting misinterpretation. "*Experience should teach us to be most on our guard to protect liberty when the government's purposes are beneficent. Men born to freedom are naturally alert to repel invasion of their liberty by evil-minded rulers. The greatest dangers to liberty lurk in insidious encroachment by men of zeal, well-meaning but without understanding*" [14].

5 Conclusions and discussion

This paper presents a review and analysis of the literature on the ethics of using publicly-available data, particularly concerning privacy. It presents the characteristics of publicly-available data and explores potential ethical issues, such as surveillance, becoming accustomed to the surveillance society, increasing access to data, combining and processing data and assuming one has nothing to hide. There is clearly much research that still needs to be done on these issues, particularly given the different perspectives on vales and ethics due to culture, religion, politics, experiences, age, gender, social status and so on.

This research comes out of the CSIR's Mobile Data Platform for Urban Mobility (MDP) work package of the Spatial Urban Dynamics 2014/2015 project and the PhD research of the first author. We would like to thank Quintin van Heerden, Peter Schmitz and Derrick Kourie for their contributions to developing this research.

References

1. Shelton, T., Poorthuis, A. & Zook, M.: Social Media and the City: Rethinking Urban Socio-Spatial Inequality Using User-Generated Geographic Information. Landscape and Urban Planning 142, 198-211 (2015).
2. Engelhardt, T.: The NSA mistakes omniscience for omnipotence. The Nation (12 Nov 2013).
3. Ioannidis, J.: Anticipating consequences of sharing raw data and code and of awarding badges for sharing, Journal of Clinical Epidemiology (2016). DOI: j.jclinepi.2015.04.015.
4. Asheim, L.: Selection and censorship: A reappraisal. Wilson Library Bulletin (R), 58 N, 180–184 (Sep 1953).
5. Asheim, L.: Not censorship but selection. Wilson Library Bulletin, 28, 63–67 (Nov 1983).
6. Channel24: A frustrated Toya Delazy leaks her own album online (14 May 2015).

7. Ilves, T.H.: Keynote Address by President Toomas Hendrik Ilves at Panel Discussion "A Secure and Free Internet", the UN Dag Hammarskjöld Library Auditorium. Permanent Representation of Estonia to the UN (29 Sep 2013).
8. Kretzschmar, L., Prinsloo, F., Prozesky, M., Rossouw, D., Sander, F., Siebrits, J., & Woermann, M.: Ethics for Accountants and Auditors. 3rd edn. Oxford University Press, Cape Town (2013).
9. Cooper, A.K., Ittmann, H.W., Stylianides, T. & Schmitz, P.M.U.: Ethical issues in tracking cellphones at an event. OMEGA, 37(6), 1063-1072 (2009).
10. Vidgen, R., Hindle, G. & Randolph, I.: Exploring the ethical implications of business analytics with a business ethics canvas. European Journal of Operational Research 281, 491–501 (2020). DOI: j.ejor.2019.04.036.
11. Yavary, A., Sajedi, H. & Saniee Abadeh, M.: Information verification in social networks based on user feedback and news agencies. Social Network Analysis and Mining, 10:2 (2020). DOI: 10.1007/s13278-019-0616-4.
12. Cooper, A.K., Schmitz, P.M.U. & Krygsman, S.C.: Tracking cellular telephones to build transport models. In: South African Transportation Conference (SATC), Pretoria, (16-19 Aug 2010).
13. Warren, S.D. & Brandeis, L.D.: The right to privacy. Harvard Law Review, 4(5) (15 Dec 1890).
14. Brandeis, L.D.: Olmstead v. United States/Opinion of the Court, 277 U.S. 438, United States Supreme Court, dissenting opinion (4 Jun 1928).
15. OAIC: What is privacy?, Office of the Australian Information Commissioner (OAIC), (2015) http://www.privacy.gov.au/aboutprivacy/what.
16. European Parliament: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). OJ L 119, 4.5.2016, p 1 (23 May 2018).
17. South Africa: Protection of Personal Information Act (Act No 4 of 2013).
18. Von Drehle, D.: The surveillance society: Secrets are so 20th century now that we have the ability to collect and store billions of pieces of data forever. TIME (1 Aug 2013).
19. Farwell, L., Richardson, D.C. and Richardson, G.M.: Brain fingerprinting field studies comparing p300-mermer and p300 brainwave responses in the detection of concealed information. Cognitive Neurodynamics, 7(4), 263–299 (2013).
20. Meijer, E.H., Ben-Shakhar, G., Verschuere, B. and Donchin, E.: A comment on Farwell (2012): brain fingerprinting: a comprehensive tutorial review of detection of concealed information with event-related brain potentials. Cognitive Neurodynamics, 7(2), 155–158 (Apr 2013).
21. Holputch, A.: Brazil's controversial plan to extricate the internet from US control. The Guardian (20 Sep 2013).
22. Cowie, J.: The new threat: Targeted internet traffic misdirection. Renesys Blog (19 Nov 2013). http://www.renesys.com/2013/11/mitm-internet-hijacking/.
23. Erickson, J.D.: The LACIE experiment in satellite aided monitoring of global crop production. In: Woodwell, G.M. (ed), The Role of Terrestrial Vegetation in the Global Carbon Cycle: Measurement by Remote Sensing, John Wiley & Sons, Ltd (1984).
24. Hudson, A.: Is small print in online contracts enforceable? BBC News: Technology (5 Jun 2013).
25. Jiow, H.J. & Lin, J.: The influence of parental factors on children's receptiveness towards mobile phone location disclosure services. First Monday, 18(1), (7 Jan 2013) doi:10.5210/fm.v18i1.4284.

26. Schweller, R.L.: The age of entropy: Why the new world order won't be orderly. Foreign Affairs (16 Jun 2014).
27. Investors.com: So Why Didn't NSA Catch The Tsarnaev Brothers? Investor's Business Daily, Inc (13 Jun 2013).
28. Leonnig, C. & O'Keefe, E.: Contractor would not have hired shooter if past brushes with law were known. TheWashington Post (17 Sep 2013).
29. Ananthaswamy, A.: Welcome to the age of the splinternet. New Scientist, (2821) (20 Jul 2011).
30. Elmer-Dewitt, P.: First nation in cyberspace. TIME International, (49) (6 Dec 1993).
31. Bentham, J.: Panopticon; or the inspection-house: Containing the idea of a new principle of construction applicable to any sort of establishment, …. Letters of Bentham (1787). Transcription and HTML by Cartome, 16 June 2001, from Bentham, Jeremy The Panopticon Writings. Ed. Miran Bozovic (London: Verso, 1995). p. 29-95.
32. Foucault, M.: 'Panopticism' from 'Discipline & punish: The birth of the prison'. Race/Ethnicity: Multidisciplinary Global Contexts, 2(1), 1–12 (Autumn 2008). English Translation by Alan Sheridan, 1977 (New York: Pantheon). Originally published in French in 1975 as Surveiller et Punir (Paris: Editions Gallimard).
33. Brunon-Ernst, A.: Introduction. In: Beyond Foucault: New perspectives on Bentham's Panopticon, Ashgate Publishing, Ltd (2012).
34. SparkNotes Editors: Sparknote on discipline and punish. SparkNotes LLC (nd). http://www.sparknotes.com/philosophy/disciplinepunish/.
35. Carr, N.: Tracking is an assault on liberty, with real dangers. The Wall Street Journal (6 Aug 2010).
36. Dobson, J.E. & Fisher, P.F.: Geoslavery. IEEE Technology and Society Magazine, pp 47–52 (Spring 2003).
37. Orwell, G.: Nineteen eighty-four. eBook No 0100021, Project Gutenberg of Australia (1949). http://gutenberg.net.au/ebooks01/0100021.txt.
38. Schmitz, P.M.U. & Cooper, A.K.: Using cellular telephones to track participants' movements to and from an event. In: South African Transportation Conference (SATC), Pretoria (11-14 July 2011).
39. Vincent, D.: Surveillance, privacy and history. History and Policy (Oct 2013). http://www.historyandpolicy.org/papers/policy-paper-151.htm.
40. Lanier, J.: Digital passivity. The New York Times (27 Nov 2013).
41. Foucault, M.: The subject and power. Critical Inquiry, 8(4), 777–795 (Summer 1982).
42. Stirling, P., Chevallier, P. and Illien, G.: Web archives for researchers: Representations, expectations and potential uses. D-Lib, 18(3/4) (2012).
43. Court of Justice of the European Union: Judgement of the Court (Grand Chamber) in Case C-131/12, Google Spain SL, Google Inc v Agencia Española de Protección de Datos (AEPD), Mario Costeja González. InfoCuria (13 May 2014).
44. Court of Justice of the European Union: Advocate General's Opinion in Case C-131/12, Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González. PRESS RELEASE No 77/13. InfoCuria (25 Jun 2013).
45. Lane, E.: Google removes 12 BBC News links in 'right to be forgotten'. BBC News: Technology (19 Aug 2014). http://www.bbc.com/news/technology-28851366.
46. Peston, R.: Why has Google cast me into oblivion? BBC News: Business (2 Jul 2014). http://www.bbc.com/news/business-28130581.
47. Scott, M.: Google reinstates some links in Europe. The New York Times (4 Jul 2014).

48. Zittrain, J.: Is the EU compelling Google to become about.me? The Future of the Internet (13 May 2014). http://blogs.law.harvard.edu/futureoftheinternet/2014/05/13/is-the-eu-compelling-google-to-become-about-me/.
49. Shein, E.: Ephemeral data. Communications of the ACM, 56(9), 20–22 (Sep 2013).
50. US Court of Appeals for the Ninth Circuit: Perfect 10, Inc v Amazon.com, Inc and A9.com Inc and Google Inc. Case F.3d, US Court of Appeals for the Ninth Circuit (2007).
51. Alexander, R.: Are search engine result figures accurate? BBC News Magazine (20 Feb 2012). http://www.bbc.co.uk/news/magazine-17068044.
52. Cooper, A.K., Byleveld, P. & Schmitz, P.M.U.: Using GIS to reconcile crime scenes with those indicated by serial criminals. In: 5th Annual International Crime Mapping Research Conference, Dallas, Texas, USA (1–4 Dec 2001).
53. Schmitz, P.M.U., Cooper, A.K., Byleveld, P. & Rossmo, D.K.: Using GIS and digital aerial photography to assist in the conviction of a serial killer. In: 4th Annual International Crime Mapping Research Conference, San Diego, California, USA (9–12 Dec 2000).
54. Gutmann, M.P. & Stern, P.C. (eds): Putting people on the map: Protecting confidentiality with linked social-spatial data. National Academies Press (2007) ISBN 978-0-309-10414-2
55. De Montjoye, Y.-A., Hidalgo, C.A., Verleysen, M. & Blondel, V.D.: Unique in the crowd: The privacy bounds of human mobility. Scientific Reports, 3(1376), 1-5 (25 Mar 2013).
56. Michalevsky, Y., Boneh, D., Schulman, A. & Nakibly, G.: PowerSpy: Location tracking using mobile device power analysis. arXiv, (1502.03182v1) (2015).
57. Siddle, J.: I know where you were last summer: London's public bike data is telling everyone where you've been. The Variable Tree (10 Apr 2014). http://vartree.blogspot.co.uk/2014/04/i-know-where-you-were-last-summer
58. Zhou, M.X., Wang, F., Zimmerman, T., Yang, H., Haber, E. & Gou, L.: Computational discovery of personal traits from social multimedia. In: 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), San Jose, CA, pp 1–6 (15–19 Jul 2013), doi: 10.1109/ICMEW.2013.6618398
59. Timberg, C.: For sale: systems that can secretly track where cellphone users go around the globe. Washington Post (24 Aug 2014).
60. Narayanan, A.: There is no such thing as anonymous online tracking. The Center for Internet and Society at Stanford Law School (28 Jul 2011). http://cyberlaw.stanford.edu/node/6701.
61. Nimmer, M.B.: The right of publicity. Law & Contemporary Problems, 19(2), 203-223 (1954).
62. Harvey, F.: To volunteer or to contribute locational information? Towards truth in labeling for crowdsourced geographic information. In: Sui DZ, Elwood S & Goodchild MF (eds), 2013, Crowdsourcing Geographic Knowledge. Springer, pp 31–42 (2013).
63. Supreme Court of the United States: Ohio v Matthew Reiner, No. 532, U.S., (19 Mar 2001).
64. Solove, D.J.: 'I've got nothing to hide' and other misunderstandings of privacy. San Diego Law Review, 44 (2007). URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=998565.
65. Big Brother Watch: Briefing note: Why communications data (metadata) matter. What are communications data? Big BrotherWatch, London, United Kingdom.
66. Collins, K.: You have more to hide in your data trail than you think. Wired UK (19 Aug 2014).
67. Phillipson, G.: Q&A: The right to privacy. BBC Religion & Ethics (14 Jun 2013).