Sustaining Legitimacy and Trust in a Data-Driven Society

Larsson, Stefan

Published in:
Ericsson Technology Review

2017

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Larsson, S. (2017). Sustaining Legitimacy and Trust in a Data-Driven Society. *Ericsson Technology Review*, *94*, 40-49.

Total number of authors:
1

Guest author

DR. STEFAN LARSSON

◆ Stefan Larsson is an associate professor at Lund University Internet Institute (LUii). He holds two Ph.D.s and an LL.M., and is an expert on digital socio-legal change, including issues of trust, consumption, traceability and privacy. He is a member of the scientific board of the Swedish Consumer Agency and recently published the book *Conceptions in the Code: How Metaphors Explain Legal Challenges in Digital Times* with Oxford University Press.

For more information about his work, visit:
http://luii.lu.se/about/stefan-larsson/

SUSTAINING legitimacy & trust IN A DATA-DRIVEN SOCIETY

Human-centric data is at the core of the digital economy and most consumer-targeted innovation. What we sometimes forget, however, is that the quantification of everyday human life that produces this data depends not only on technological capabilities, but also on social norms and user values.

■ **TRUST IS A VITAL** determining factor influencing users' decisions to adopt innovations and sign up for new services – particularly those that they know will generate data for the service provider. While user trust is heavily based on their perception of the technological security of a solution or service, it is also fundamentally dependent on social norms and values such as privacy, legitimacy and perceived fairness in the collection and handling of individual information. The long-term success of the digital economy is dependent on consistently high levels of both technological and sociological trust among users. In light of this, it is of utmost importance that service providers consider the implications of social norms and user values in the service design process.

Human-centric big data

A large proportion of ICT innovation today is driven by the collection and analysis of human-centric data – a key component of the big data phenomenon. In some cases, human-centric data is collected by a company from the users of its current services. In other cases, a company may have purchased the data from another company to gain a better understanding of a new target group, for example. No matter how it is sourced, data is collected, analyzed and traded on a continuous basis, acting as a backbone for wide-ranging products and services: from health to consumer goods and services to urban planning.

The implications of this trend extend far beyond mere digitalization in terms of communication and infrastructure. Several scholars have argued that the growing strength of social networks is causing society to become not only "digitized" but increasingly "datafied" – with profound effects on how we read, write, consume, use credit, pay taxes and educate ourselves [1, 2].

This large-scale quantification of human activities has occurred within a very short period of time. Just a few years ago, it was much more difficult to gather human-centric data and use it for service development or commodification. But now, whenever we use the internet or carry a smartphone that is connected to it, we are tracked, logged, analyzed and predicted in a variety of ways: by way of web cookies, search engines, social media, e-mail and online purchases, as well as various types of sensors (including RFID tags and GPS-enabled devices such as cameras, smartphones and wearables). Offline purchase history is another useful resource, which can be administered through loyalty cards and club memberships, for example.

All of this information relating to our activities is not only used by the organizations that collect it; it is also exchanged by numerous commercial and governmental players for a whole variety of reasons. Beside this, there are companies known as data brokers that specialize in collecting and trading consumer data that is often at least partly collected from public sources. Such data collection and trading activities rarely involve a human observer who actually monitors the data points. They rather tend to be handled by an automated, quantitative and ubiquitous storage system built into the infrastructure – in the widest sense of the word – itself [cf. 3].

Some social scientists claim that this trend represents one of the most far-reaching social changes of the past 50 years [cf. 4]. As a result, these data-driven and technology-mediated practices are increasingly gaining the attention of scholars in various disciplines, particularly as they relate to privacy, but also in a variety of critical perspectives on transparency and algorithmic accountability [5, 6], big data ethics [7], behavioral and traditional discrimination [8, 9] or other consequences of a data-driven "platform society" [10].

Web cookies and the black box society

Web cookies are among the tools being used by companies such as Google, Facebook and traditional media houses to create extensive data retention infrastructures. The 2015 update of the Web Privacy Census revealed that a user who visits the world's 100 most popular websites receives more than 6,000 web cookies, which are stored on their computer [11]. Furthermore, it found that Google tracking infrastructure is on 92 of the top 100 most popular websites and on 923 of the top 1,000 websites, which contributes to making Google the world's most powerful information manager, with a central place in the modern information economy.

Similarly, a 2015 study by the Norwegian data protection authority Datatilsynet showed which parties were present when visiting the front page of six Norwegian newspapers [12]. The report noted that between 100 and 200 web cookies were placed on any computer being used to visit these home pages, that information about the visitor's IP address was sent to 356 servers, and that an average of 46 third parties were "present" during each visit. However, none of the six newspapers provided their audience with any information relating to the presence of this large selection of third-party companies.

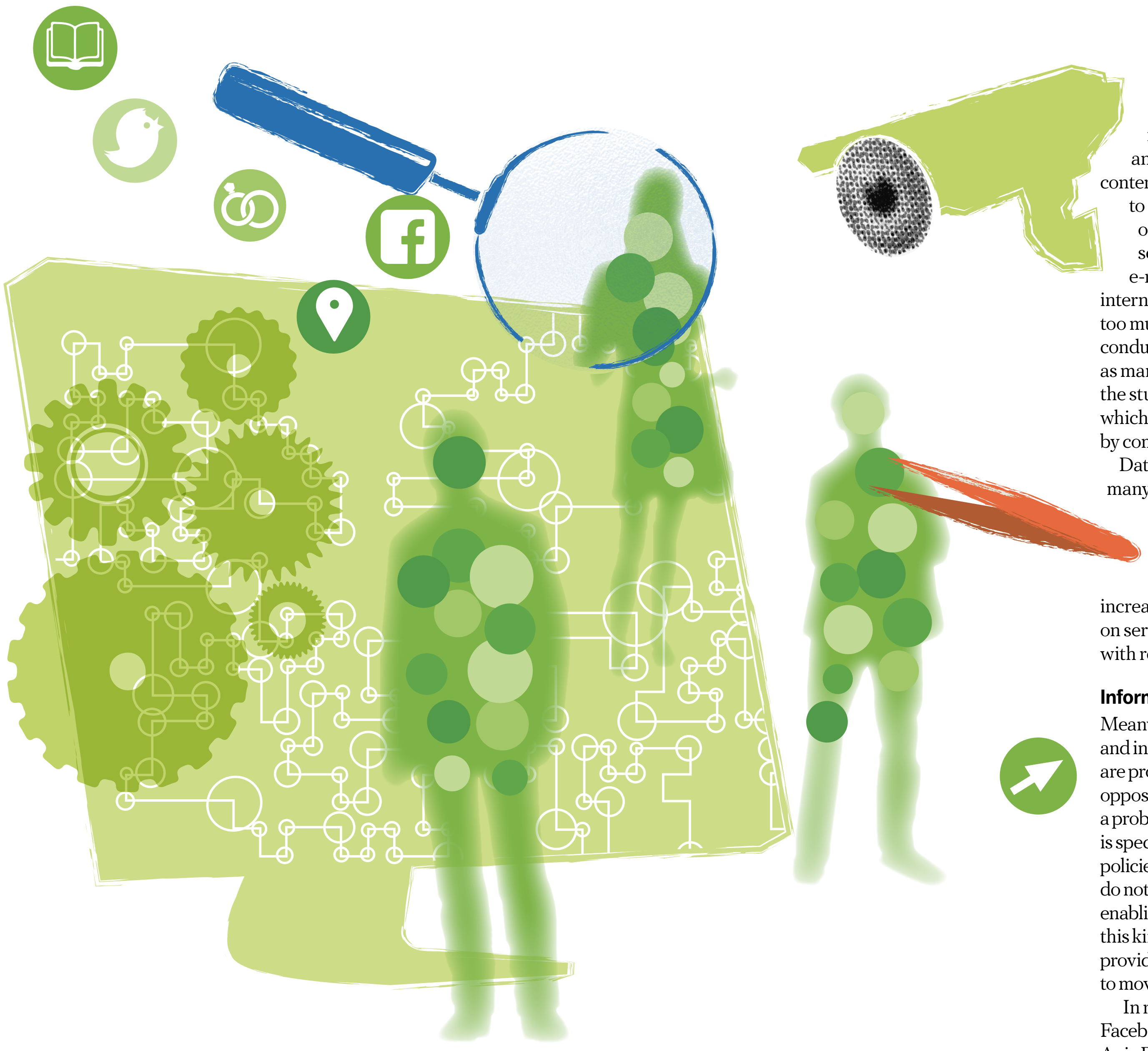
The use of web cookies in this manner contributes to the creation of what has been dubbed the "black box society" [13], where users are unable to make informed decisions when choosing services. Any attempt to find the services that are the most privacy friendly is doomed to fail because users are kept largely in the dark.

While advertising companies are the key players in this arena, theirs is far from the only segment that sees the benefits of individually targeted data-gathering practices. The ongoing introduction of innovative analytical methods adds to the importance of the data, including the shift from descriptive to predictive analytics [14].

Growing concerns over lack of control

Several surveys carried out in recent years have revealed that users are becoming increasingly concerned about their lack of control over the use

❝ SEVERAL SURVEYS CARRIED OUT IN RECENT YEARS HAVE REVEALED THAT USERS ARE BECOMING INCREASINGLY CONCERNED ABOUT THEIR LACK OF CONTROL OVER THE USE AND DISSEMINATION OF THEIR PERSONAL DATA ❞



and dissemination of their personal data. They are particularly worried about having no control over their internet-generated personal data, and the possibility of it being used in ways other than those they originally intended when sharing it [15, 16]. Many people are concerned about the capability of third parties such as advertisers and other commercial entities to access their personal information [16, 17, 18, 19].

A clear majority of internet and online platform users in the European Commission's Special Eurobarometer 2016 expressed their discomfort over the fact that online platforms use information about their internet activities and personal data to tailor advertisements or content to their interests [20]. Further, according to the EU Commission in 2015, only 22 percent of Europeans fully trust companies such as search engines, social networking sites and e-mail services, and as many as 72 percent of internet users are worried about being asked for too much personal data online [21]. In a survey conducted by the Pew Research Center in 2014, as many as 91 percent of US users who took part in the study felt they had lost control over the ways in which their personal details are collected and used by companies [16].

Data collection and handling is clearly fueling many users' growing sense of distrust in service and goods providers. This is naturally a great cause for concern since access to user data is a key enabler of the digital economy. At a certain point, the users' increasing unease could have a damaging effect on service usage levels, and serious repercussions with respect to the digital economy as a whole.

Information overload

Meanwhile, just as the lack of consumer control, and in a sense, the shortage of available information, are problematic, there are indications that the exact opposite – information overload – is also presenting a problem. The information overload in question is specifically related to user agreements, privacy policies and cookie usage. Online user agreements do not appear to be particularly effective in terms of enabling informed user choices. Critics argue that this kind of "privacy self-management" does not provide meaningful control and that there is a need to move beyond relying too heavily on it [22].

In relation to a study on consent practices on Facebook, media scholar and digital sociologist Anja Bechmann posits that "the consent culture of the internet has turned into a blind non-informed consent culture" [23, p. 21]. User agreements often constitute little more than an alibi for providing data-driven businesses with access to user data. The validity of this kind of agreement is consequently questionable.

The trouble with these agreements is that they tend to be too long, too numerous and too obscure. The result is that most users don't read them carefully and are therefore not fully aware of what they are agreeing to when they sign them. For example, a study that tracked the internet browsing behavior of 48,000 monthly visitors to the websites of 90 online software companies found that only one or two of every 1,000 retail software shoppers accessed the license agreement, and that most of those who did access it read no more than a small portion. The conclusion in that study was that the limiting factor in becoming informed thus seemed not to be the cost of accessing license terms but reading and comprehending them [24, cf. 25]. Arguably, the sheer amount of lengthy license agreements that even an average user of digital services agrees to constitutes a sort of information overload. For example, Norway's consumer ombudsman Forbrukerrådet recently conducted a study that involved reading the terms and conditions of all the apps on an average smartphone. Reading them was found to take 31 hours and 49 minutes [26].

Media researcher Helen Nissenbaum has pointed out that the obscurity of the agreements may serve a purpose: if they were written more clearly, they would likely be far less readily accepted [27]. In a recent study, the privacy policies of 75 companies that track behavior in digital contexts were reviewed, and the researchers found that many of them lacked important consumer-relevant management information, particularly with respect to the collection and use of sensitive information, the tracking of personally identifiable data and companies' relationships to third parties [28]. In the short term, a fuzzy and extensive privacy policy appears to be a helpful tool in the data-gathering race. But will there be a price to pay in the long run?

The privacy paradox and acceptance creep

In many cases, there is a significant gap between a service provider's commercial data practices and the normative

> “PERCEPTIONS OF PRIVACY AND SOCIAL NORMS RELATING TO COMMERCIAL USE OF INDIVIDUAL DATA CHANGE DYNAMICALLY OVER TIME DUE TO SOCIO-TECHNOLOGICAL SHIFTS IN GENERAL, AND IMPROVED SERVICES IN PARTICULAR”

preferences of many – or even most – of its users. Yet research shows that many users often continue to use services that can be very intrusive, while at the same time stating that they are concerned about data being collected when they use products and services online [cf. 23]. Other studies demonstrate that many individuals have not made any major changes to their data sharing or privacy practices in recent years, despite their concerns regarding online data collection [17, 29, 30, 31]. In our behavior, we tend to “accept the cost of free,” as noted by competition law scholars Ariel Ezrachi and Maurice Stucke [8, p. 28].

US consumption researchers have put this “privacy paradox” down to consumers’ sense of resignation toward the use of their personal data [32]. In the case of loyalty cards, studies show that although consumers do not necessarily feel satisfied with receiving discounts as a trade-off for sharing their personal data, they feel resigned about the situation rather than driven to address the imbalance.

Are these all signs that we are experiencing a phenomenon that legal scholars Mark Burdon and Paul Harpur [33] call “acceptance creep,” with massive data collection practices becoming normalized among users? If so, does the acceptance creep merely point to a sense of resignation (too many choices, too much information – resistance is futile) or to the beginning of a fundamental shift in social norms (perceptions) regarding data and privacy?

The answer likely contains a little bit of both. Perceptions of privacy and social norms relating to commercial use of individual data change dynamically over time due to socio-technological shifts in general, and improved services in particular. But the current gap between the stated norms of users and the data practices of service providers is very clear. A great deal of the commercial data collection and handling that is taking place at present is simply not perceived as legitimate. Figuring out how to handle users’ normative and behavioral preferences and navigating the “non-informed” consent culture is a major challenge for service designers in a data-driven digital economy.

Ethical implications of information asymmetry

The emergence of big data has added to the information asymmetry between customers and the companies in the insurance, airline and hotel industries and other traditional markets – an effect that is further amplified by the advent of predictive analytics [cf. 8]. This raises several questions about service development and design in terms of how the more qualitative aspects of humanity might be incorporated into all of this quantification. The first question relates to balancing powers on the markets, which in most cases would mean empowering the consumers who are often in the dark with regard to how their data is being collected, analyzed and traded. One way of doing this would be to increase transparency about data practices; another would be to redesign the legal and structural protective measures to better protect weaker parties that have provided “non-informed consent” from being taken advantage of by service providers.

A somewhat more complex question that needs to be addressed is the extent to which users’ values and cultures should be considered when designing large-scale automated systems and algorithms. This is related to the ethical and moral questions that may arise as an outcome of quantification and automation of a particular kind. Concerns like this have only begun to be conceptualized and discussed – one example being a recent report from the committees of the IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems [34].

We can already see a growing tendency among market players such as insurance brokers, money lenders and health institutions to base interest rates, insurance costs and payment plans on detailed, big-data-based analyses of individuals. This could also concern predictive analytics of future health, income and life expectancy.

What are the potential risks and repercussions of this kind of development from an ethical and normative point of view? More specifically, how should we understand and govern complex (and often proprietary) algorithms and machine-learning processes that may produce troublesome consequences from a social, legal, democratic or other perspective? These are questions that both the public and private sectors need to address urgently.

Commercial practices and the lagging law

One of the key challenges met when regulating the use of human-centric data is that the use of such information has already become so integral to innovation at a time when both lawmakers and private individuals are still largely unaware of how it is collected and used. From a legal point of view, the challenge is arguably largely the result of a lack of knowledge of growing data practices and their outcomes, but is also of a conceptual kind: how should new practices and phenomena be understood and governed? Law is inevitably path dependent in that it is reliant on past notions or past social and technological conditions when regulating contemporary challenges. The result, according to emerging socio-legal research in the field, is a sort of path-dependent renegotiation of traditional concepts for the regulation of new phenomena [35]. For example, should Facebook be liable for content in that same way as a traditional news outlet when mediating news for its 1.79 billion monthly active users (as of the third quarter of 2016)? Should Uber be regarded as a taxi company and an employer, and be taxed accordingly in each of the more than 60 countries it operates in?

Given that contemporary digital innovation is often disruptive (creating new markets and value networks, and displacing established firms, products and partnerships), the development of new services and products tends to be carried out iteratively. In light of all of this, the fact that the law is lagging is therefore not surprising or strange. Nonetheless, it is vital that we continuously strive to close the gap – particularly in the face of new conceptual dilemmas that involve data-driven innovation, legitimacy and trust.

Conclusion

From a legal point of view, I think regulators must develop a more critical perspective and a better understanding of how to manage data-driven and algorithm-controlled processes as well as data analyses. They must continuously improve their ability to recognize when consumers need protection and empowerment, and strive for transparency with regard to how new technologies work and what kinds of regulations we need to ensure that future developments are in users' best interest.

Ultimately, the continued success and future development of the digital economy will depend on our ability to strike a balance between the interests of individuals, commercial players and governments when it comes to data collection and usage. While regulation in the form of laws such as the Swedish Personal Data Act (Personuppgiftslagen or PuL) and the EU's General Data Protection Regulation (GDPR) will continue to play an important role, the pace of technological development is likely to continue to leave lawmakers playing catch-up.

It is therefore crucial for the private sector to take a proactive approach to addressing normative and ethical questions as part of the service design and development process. Otherwise, there is a significant risk that consumers' trust in digital services will decline in the mid to long term. A low level of trust in new features, services and devices could substantially reduce their potential scalability, and consequently have a negative impact on the digital economy as a whole. ⊛

References

1. **Kitchin, R.** *(2014) The Data Revolution. Big data, open data, data infrastructures & their consequences,* **SAGE**
2. **Mayer-Schönberger & Cukier (2013)** *Big Data – A Revolution That Will Transform How We Live, Work, and Think, Boston and New York:* **Eamon Dolan/ Houghton Mifflin Harcourt**
3. **Andrejevic, M. (2013)** *Infoglut. How too Much Information is Changing the Way We Think and Know. New York, NY:* **Routledge**
4. **Rule, J.B. (2012) "Needs" for surveillance and the movement to protect privacy. In K. Ball, K. D. Haggerty & D. Lyon (eds.)** *Routledge handbook of surveillance studies* **(pp. 64-71). Abingdon, Oxon: Routledge**
5. **Rosenblat, A., Kneese, T. & Boyd, D. (2014) "Algorithmic Accountability." A workshop primer produced for The Social, Cultural & Ethical Dimensions of "Big Data" March 17, 2014, NY**
6. **Kitchin, R. & Lauriault, T.P. (2014) Towards critical data studies: Charting and unpacking data assemblages and their work,** *The Programmable City Working Paper 2*
7. **Richards, N.M. & King, J.H. (2014) Big Data Ethics, 49 Wake Forest Law Review, 393-432**
8. **Ezrachi, A. & Stucke, M.E. (2016)** *Virtual Competition. The Promise and Perils of the Algorithm-Driven Economy.* **Harvard University Press**
9. **Datta, A., Tschantz, M.C., Datta, A. (2015) Automated Experiments on Ad Privacy Settings – A Tale of Opacity, Choice, and Discrimination.** *Proceedings on Privacy Enhancing Technologies***. 1: 92–112, DOI: 10.1515/ popets-2015-0007**
10. **Andersson Schwarz, J. (2016) "Platform logic: The need for an interdisciplinary approach to the platform-based economy", paper presented at** *IPP2016***: The Platform Society, Oxford Internet Institute**
11. **Altaweel, I., Good, N. & Hoofnagle, C. (2015) Web privacy census.** *Technology Science*
12. **Datatilsynet (2015) The Great Data Race.** *How commercial utilization of personal data challenges privacy*
13. **Pasquale, F. (2015)** *The Black Box Society. The Secret Algorithms That Control Money and Information***, Harvard University Press**
14. **Siegel, E. (2016) Predictive Analytics:** *The Power to Predict Who Will Click, Buy, or Die.* **Wiley**
15. **Lilley, S., Grodzinsky, F.S. & Gumbus, A. (2012)** *Revealing the commercialized and compliant Facebook user.* **Journal of information, communication and ethics in society, 10(2): 82-92**
16. **Pew (2014)** *Public Perceptions of Privacy and Security in the Post-Snowden Era.* **Pew Research Center**
17. **Findahl, O. (2014)** *Svenskarna och Internet 2014.* **Göteborg: .SE**
18. **Kshetri, N. (2014)** *Big data's impact on privacy, security and consumer welfare.* **Telecommunications Policy, 38(11)**
19. **Narayanaswamy, R. & McGrath, L. (2014)** *A Holistic Study of Privacy in Social Networking Sites***, Academy of Information and Management Sciences Journal, 17(1): 71-85**
20. **Special Eurobarometer 447 (2016) Online Platforms**
21. **COM (2015) 192 final.** *A Digital Single Market Strategy for Europe*
22. **Solove, D.J. (2013)** *Privacy Self-Management and the Consent Dilemma.* **126 Harvard Law Review 1880-1903**
23. **Bechmann, A. (2014)** *Non-informed consent cultures: Privacy policies and app contracts on Facebook.* **Journal of Media Business Studies, 11(1): 21-38**
24. **Bakos, Y., Marotta-Wurgler, F. & Trossen, D.R. (2014)** *Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts.* **Journal of Legal Studies, 43(1): 9-40**
25. **McDonald, A.M. & Cranor, L.F. (2008)** *The Cost of Reading Privacy Policies.* **Journal of Law and Policy for the Information Society**
26. **Forbrukerrådet (24 May 2016)** *"250,000 words of app terms and conditions"* **http://www.forbrukerradet.no/ side/250000-words-of-app-terms-and-conditions/**
27. **Nissenbaum, H. (2011)** *A contextual approach to privacy online,* **140 DAEDALUS 32-48**
28. **Cranor, L.F., Hoke, C., Leon, P.G. & Au, A. (2014)** *Are They Worth Reading? An In-Depth Analysis of Online Advertising Companies' Privacy Policies,* **TPRC Conference Paper**
29. **Christensen, M. and Jansson, A. (2015)** *Complicit surveillance, interveillance, and the question of cosmopolitanism: Toward a phenomenological understanding of mediatization.* **New Media & Society, 17(9): 1473-1491**
30. **Light, B. & McGrath, K. (2010)** *Ethics and Social Networking Sites: a disclosive analysis of Facebook.* **Information, technology and people, 23(4): 290-311**
31. **Martin, S., Rainie, L., & Madden, M. (2015)** *Americans Privacy Strategies Post-Snowden.* **Pew Research Center**
32. **Turow, J., Hennessy, M., Draper, N. (2015)** *The Tradeoff Fallacy: How Marketers Are Misrepresenting American Consumers and Opening Them Up to Exploitation.* **Research report, Annenberg School of Communication, University of Pennsylvania**
33. **Burdon, M. & Harpur, P. (2014)** *Re-conceptualizing Privacy and Discrimination in an Age of Talent Analytics,* **University of New South Wales Law Journal 37(2): 679-712**
34. **The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems (2016)** *Ethically Aligned Design: A Vision For Prioritizing Wellbeing With Artificial Intelligence And Autonomous Systems***, Version 1. IEEE. http://standards. ieee.org/develop/indconn/ec/ autonomous_systems.html**
35. **Larsson, S. (2017)** *Conceptions in the Code: How Metaphors Explain Legal Challenges in Digital Times.* **Oxford University Press.**