Philos. Technol. (2013) 26:357–379
DOI 10.1007/s13347-013-0104-0

SPECIAL ISSUE

Balance or Trade-off? Online Security Technologies and Fundamental Rights

Mireille Hildebrandt

Received: 27 January 2013 / Accepted: 29 April 2013 / Published online: 21 May 2013

Abstract In this contribution, I will argue that the image of a balance is often used to defend the idea of a trade-off. To understand the drawbacks of this line of thought, I will explore the relationship between online security technologies and fundamental rights, notably privacy, nondiscrimination, freedom of speech and due process. After discriminating between three types of online security technologies, I will trace the reconfiguration of the notion of privacy in the era of smart environments. This will lead to an inquiry into the metaphor of the scale, building on the triple test regarding the justification of the limitation of fundamental rights such as privacy. The conclusion will be that in the case of a trade-off, infringing measures will have to be balanced by effective safeguards. No trade-off without balance.

Keywords Online security technologies · Security · Safety · Privacy · Monitoring · Filtering · Confidentiality · Integrity · Availability · Cybercrime · Balance · Trade-off

1 Introduction

We are often told that we must sacrifice some of our liberties to gain security. Though this is usually explained as a matter of balancing liberty and security, it refers to a trade-off, not a balance. The idea of balancing would imply that whenever we increase the employment of security measures that violate our liberties, this warrants extra safeguards to regain the balance. In this contribution, I will focus on the

M. Hildebrandt (✉)
Institute of Computing and Information Sciences (iCIS),
Radboud University Nijmegen, Nijmegen, Netherlands
e-mail: hildebrandt@frg.eur.nl

M. Hildebrandt
Erasmus School of Law (ESL), Rotterdam, Netherlands

M. Hildebrandt
Law Science Technology and Society (LSTS), Vrije Universiteit Brussel, Brussel, Belgium

infringements that online security technologies (OST) generate for our fundamental rights and investigate what is meant when the image of the scale is invoked: must we indeed trade some of our liberty to gain security, or do we need more safeguards as countervailing measures against violation of our civil rights? My conclusion will be that it is very well possible that in specific instances, security measures are necessary to prevent and resolve online security problems, but that we cannot assume that OSTs necessarily improve our security. This will require empirical evidence, which is often hard to find. Furthermore, in the cases where OSTs are necessary, any trade-off should be balanced with specific and effective safeguards regarding our fundamental rights.

In Section 2, I will discuss online security and OSTs to prevent a Babylonian speech confusion around concepts like security and safety, threats and vulnerabilities, cyber security and cyber crime and finally between security measures and security technologies. Though all these terms are interrelated, they require explanation, also concerning their overlap, precisely because scholars and experts from different professional or disciplinary backgrounds may have alternative understandings of what is meant. Apart from these terminological issues, I find that lawyers and computer scientists may have a very different comprehension of what constitutes a threat or vulnerability and, for instance, to what extent technical solutions produce new vulnerabilities. The third reason to pay substantial attention to which OSTs are operational for what purpose is to make sure that security threats are not seen as merely a social construction, though they are certainly constructed. With Latour I dare say that their constructive ontology makes them even more real.[1] Coming to terms with freedom infringements requires that we take serious the security issues generated by the cocktail of an exponential increase (1) in remote control, (2) scale and (3) speed, (4) hyper-connectivity and (5) automation.

In Section 3, I will discuss the relevant fundamental rights and briefly explain how they are impacted by the transformation of our information and communication infrastructure, and in particular by OSTs that try to cope with transnational, automated, large-scale, high-speed, and hyper-connected security issues. Instead of merely discussing privacy in the common sense meaning of a right to minimal disclosure, I explain that privacy is a more complex right, best described as the freedom from unreasonable constraints on identity construction, while taking into account that a number of other fundamental rights are at stake, notably data protection, non-discrimination, due process and free speech. Instead of lamenting the technological transition of our life world, I propose a more interactive approach to privacy, and the need for counter infringement measures.

In Section 4, I discuss the image of the scale, as dissected by legal philosopher Jeremy Waldron, taking six caveats from his analysis that are relevant for the notion of a trade-off and/or balance between OSTs and fundamental rights. Based on these caveats, I explain how current human rights law within the context of the European Convention of Human Rights organises the limitation of fundamental rights,

[1] Surveillance studies generally focus on privacy and discrimination issues generated by surveillance techniques and technologies (see, e.g. the Surveillance Studies Network, at http://www.surveillance-studies.net/?page_id=119). This article starts with a brief assessment of security and other issues that may call for OSTs, before investigating their implications for our fundamental rights. On Latour's 'realistic realism', which I would term 'constructivist realism', see Stalder (2000).

elaborating on the so-called triple test that OSTs that infringe our rights and liberties must meet. This test, requiring (1) a legitimate aim, (2) necessity and proportionality and (3) a legal ground that incorporates specificity, foreseeability as well as adequate and effective safeguards, has been developed and adapted in the course of decennia of human rights case law, in confrontation with a myriad of concrete cases, integrating the acuity and discernment of generations of excellent lawyers from around Europe. It takes into account under what conditions a trade-off may be necessary and which requirements must be met to achieve the right balance. Though the European Convention of Human Rights applies primarily to governmental interventions, its practical wisdom should inspire the choice, the design and the employment of OSTs, notably the integration of effective technical and organisational counter infringement measures.

Section 5 sums up the conclusions.

2 Online Security and OSTs

Our common sense easily conflates and confuses safety and security. In specialist discourse, safety is usually understood to refer to protection against physical or other harm, whereas security is taken to denote the prevention of or the resilience against deliberate attacks (e.g. Schneier 2006, p. 12). The concepts overlap. Safety can be increased if attacks can be prevented or withstood, but safety can also be impaired when no attacks are at stake (natural disaster, harmful side-effects of human action and unintentional negligent causing of harm). Meanwhile, the concept of security is often used to refer to the reduction of foreseeable harm (thus hopefully increasing safety), and related to resilience against crime (Zedner 2009). However, in a more technical sense it concerns the confidentiality, integrity or availability of data, computing systems or service provision which does not necessarily involve safety (Leeuw and Bergstra 2007).[2] Safety is thus both more and less than security and vice versa. It is important at this point to clarify that increased security in the technical sense does not imply increased safety, both because safety hazards exist beyond intentionally caused harm and because technical security sees to a number of issues that have no direct link with harm.

In this contribution, I want to investigate how the legal framework of fundamental rights relates to issues of online security. My interest is not restricted to an analysis of positive law but focused on the purposiveness, justice and legal certainty of law in a constitutional democracy. The German legal scholar and legal philosopher Gustav Radbruch has coined these three dimensions of the law as antinomies; they should direct the operations of the law even though they do not always work in concert (Leawoods 2000; Radbruch 1950). This requires a balancing act. For instance, if the *purpose* of the law is to increase the safety of citizens or the availability of online services, while *justice* requires equal respect and concern for all, this implies that safety and security are distributed in a fair way and requires that infringements of rights and freedoms are complemented by adequate safeguards; finally *legal certainty*

[2] If the service of a Website selling shoes suffers an attack that causes disruption in its services, this impacts its availability and may even cause financial damages. But it does not impair anybody's safety.

means that citizens can base their own actions on the expectation that legal norms will be effective and will be interpreted with a view to the integrity of the legal system as a whole. In that sense legal certainty relates to foreseeability and trust, and thus sustains the instrumentality of the purposiveness and the justice of the law: without legal certainty they will not flourish.

In relation to online security, the construction of the legal framework determines the incentive structure for public and private investment in security measures.[3] This incentive structure depends on legal certainty, defined as the integrity and foreseeability of legal effects. To generate trust and innovation a society depends on such integrity and foreseeability, notably when it comes to rights and liberties such as freedom from unwarranted interference, freedom to contract, freedom to participate in political decision making, civil and criminal liability, rights to specific public goods and the right to an effective legal remedy when private or public rights and violated. To a large extent, legal certainty eventually depends on the monopoly of violence of the modern state, which sustains the internal and external sovereignty that legitimates the social contract. The security of citizens against both internal and external violence (crime and war) is in fact constitutive of the modern state; without such security the state is neither feasible nor sustainable.[4] The problem of failed states is precisely that the state does not protect its citizens against violent attack and thus forces them to take up arms to defend themselves or to be subjected to the terror of military factions or mafia rule.

2.1 Online Security: Threats and Vulnerabilities

Online security in the technical sense can be defined in terms of threats (e.g. attack toolkits) and in terms of vulnerabilities (e.g. buffer overflow and SQL protocols). Vulnerabilities can be exploited by what is called malicious code (e.g. a virus, worm and Trojan) that is capable of destroying data or running destructive or intrusive programs, for instance stealing sensitive information (of a person or a business). Other security issues concern spam that may be sent by remotely controlled spam zombies and fraud that may be effected by phishing hosts or bot-infected computers. In its Internet Security Threat Report over 2011 (Symantec, Internet Security Threat Report 2011), Symantec reports the blocking of over 5.5 billion malware attacks, an 81 % increase over 2010; an increase of 36 % of web based attacks with over 4.500 new attacks each day; an increase of 41 % (403 million) new variants of malware compared to 2010; exposure of 232 million identities; an average of 82 targeted attacks per day. They indicate that mobile threats are increasingly collecting data, tracking users and sending premium text messages and conclude that one is more

[3] For instance, product liability for security vulnerabilities in software would create an entirely different incentive structure than the current one. At a more basic level, tort law and criminal liability—including the applicable law of evidence and burden of proof—determine who will invest up to what level in security measures.

[4] For example, Piret (2008) on the historical significance of the concept of sovereignty, articulated as a critique of Hannah Arendt's critique of sovereignty. Though I do not necessarily agree with Piret on all accounts, I believe we should acknowledge that the enforcement of safety, security and human rights protection to a large extent still depends on the monopoly of violence within sovereign states (see also Hildebrandt (2013)).

likely to be infected by malware placed on a legitimate web site than one created by a hacker. It is important to note that terms like 'attack' or 'incident' can group together events of great diversity, some or even many of which may not really be threatening at all (easily identifiable phishing attempts, readily detected viruses or inadequate attempts to log into a system). By adding these to the statistics alarmist impressions are created that do not help to find any kind of balance between measures to achieve online security and fundamental rights (e.g. Sommer and Brown 2011, p. 7).

As indicated above, security in the technical sense is usually defined in terms of confidentiality, integrity and availability (CIA) (Leeuw and Bergstra 2007) and builds on the assumption of attackers trying to violate either one. Confidentiality usually concerns data and nourishes the numerous narratives of Alice who want to send a confidential message to Bob, while Eve tries to figure out the content of the message. Eve is clearly the attacker here; she may be a person, a computer system, a Webbot. Integrity may concern either data or computer systems and here the attack aims to manipulate data or systems, which renders them incorrect or incomplete. Availability concerns a computing system or service whose resilience is broken down resulting in a loss of functionality. In addition, security also concerns authentication or identification underlying access control; attackers will for instance try to gain access to online services by spoofing (false identification) or phishing (collecting credentials for authentication). For this contribution, it is important to discriminate between the security of individual transactions or communications, that of public or private service providers and the security of critical infrastructure. The seriousness of a threat or vulnerability will depend on the expected frequency of the violation of confidentiality, integrity and/or availability and its impact. Ranking the impact will depend on the amount of individuals who will be affected, the distribution of the impact, the invasiveness, duration and secondary effects. Evidently, attacks on the availability of critical infrastructure, such as telecom providers, banks, the Smart Grid, the Internet, public transport, will impact a large number of individuals and may have a profound effect on their lives.

Threats to online security in the technical sense are not equivalent with cybercrime. In a broader sense online security may refer to resilience against cybercrime, which entails criminal offences committed with, on or against interconnected computing systems (Brenner 2007a). The Cybercrime Convention (Watney 2012) obligates states to criminalize threats to CIA, notably illegal access, illegal interception, unlawful data interference, system interference, misuse of devices (including passwords); to criminalize computer-related offences, notably computer-related forgery and fraud; to criminalize content-related offences, notably child pornography; and offences related to the infringements of copyright. Cybercrime may concern existing crimes, such as fraud, child pornography or copyright infringement that do not threaten online security in the technical sense, though the emergence of cyberspace changes the scope, speed and distance involved in these crimes, justifying special treatment. As will be discussed in the next section, measures to fight cybercrime may in fact threaten online security.

Online security is implicated in the notion of cyberwar. Though not the subject of this paper, due to the issue of extraterritorial jurisdiction-to-enforce and the blurring of internal and external sovereignty, this topic looms in the background (Hildebrandt 2013). The Stuxnet, the Flame and the Wiper virus have demonstrated that states do

engage in unilateral attacks on specific infrastructure outside their own territory without declaring war, thus seemingly transposing Grotius' *Mare Liberum* to the realm of cyberspace. One could speak of a *Cyberspace Liberum* in which no monopoly of violence is in effect and no social contract can be assumed (Hildebrandt 2013). This would require serious reflection on what kind of natural law rules in cyberspace, taking leave of unwarranted claims of cyber utopism and the dangerous cynicism of Real Politik. Grotius was neither a utopian nor a Realpolitiker.

2.2 Online Security Technologies

To evaluate the balance between online security and civil rights, we need to think in terms of online security measures on the one hand, that should promote the benefit of online security, and the costs of infringements of civil rights on the other. Note that the measures themselves are not a benefit; whether they achieve a benefit remains to be seen and this is ultimately an empirical as well as an evaluative question. Also, both the potential benefits and the costs will involve distribution issues: who share the costs and who share the benefits? From the perspective of philosophy of technology, it is interesting to narrow the analysis down to OSTs: how do different types of technologies impact fundamental rights? Building on the idea that technology is neither good nor bad but never neutral, I will therefore investigate what types of OSTs are currently used or proposed to achieve a reasonable level of online security.

OSTs can be divided in, first, technologies to ensure confidentiality of communication or rightful authentication online, second, technologies to detect and counter threats and vulnerabilities online, and third, technologies used to detect and counter cybercrime such as forgery, fraud, child pornography and copyright violations committed in cyberspace. The first type of OSTs, that are meant to ensure confidentiality, usually involve encryption and this implies that they also afford what some call data privacy or data confidentiality; to some extent they also protect against the profiling of content. The human rights of privacy and data protection are often defined in terms of secrecy or data minimisation; though this covers only part of these rights it seems that encryption does not violate fundamental rights (Gürses et al. 2006; Section 2). However, encryption often depends on trusted third parties for key management and certification. To the extent that trust is intransitive and does not scale, this implies that encryption generates new vulnerabilities that require further OSTs to detect fraudulent third parties or attacks against trusted platforms. Similarly, OSTs which ensure authentication, may seem neutral as to human rights. This conclusion is again premature. Authentication allows access control and though this may be a benefit for individual citizens (to the extent that it prevents unauthorized access to personal data by others) it all depends on who is in control, how much transparency is built-in and what remedies individuals have against illegitimate access to their data. Authentication technologies allow to exclude people, they afford walled gardens on the Web and within a secured environment these techniques may also afford the monitoring of individuals by means of re-recognition (e.g. Bennett and Lyon 2008; Gandy 2000).

The second type of OSTs, that are meant to detect and counter online threats and online vulnerabilities, can be divided in monitoring, filtering and blocking technologies.[5] Monitoring requires analysing Internet traffic, which means that volume, address, routing or even content of the packets that co-constitute the Internet are inspected to detect various types of malware or illegal content. Filtering removes or diverts Internet traffic, e.g. viruses, malicious users, child pornography, advertisements, or hate speech. Filtering is automated and based on specific rules that may be either over inclusive or under inclusive (filtering more or filtering less then intended). Note that child pornography, advertisements or hate speech do not threaten online security, and advertisements are not necessarily illegal, but filtering that type of content is performed by using the same technology that protects against malware. Blocking entails preventing a specific IP address or set of IP addresses from uploading and/or downloading content to or from the Internet, for instance to prevent illegal filesharing, to prevent malious attacks or to prevent a computing system from controlling other computers by means of a botnet. Monitoring is connected with surveillance, and filtering and blocking is often considered to be a violation of so-called Net Neutrality.

The third type of OSTs, that are meant to detect and counter cybercrime, can be subdivided in technologies that enable the execution of justice authority competences, such as those imposed on states in the Cybercrime Convention: expedited preservation of stored computer data, production orders, search and seizure of stored computer data, real-time collection of computer data, and interception of content data. The most relevant technologies in relation to online security are those that enable remote access to and remote control of computing systems (hacking), which includes disabling, disconnecting, or even destroying data and software running on a remote computing system (server). Obviously, these technologies are precisely those that threaten the CIA of data and software and related services—thus compromising online security in the technical sense. This means that the use of these technologies by law enforcement authorities falls within the scope of criminal offences, requiring specific legal competence to provide for a justification.

Unsurprisingly, OSTs generate new online security risks as well as violations of fundamental rights: encryption requires trust that can be violated; authentication allows exclusion and may enable monitoring; monitoring is connected with surveillance which threatens confidentiality and privacy, and surveillance also links with social sorting which threatens equal treatment; and finally filtering and blocking is often considered a violation of so-called Net Neutrality that should guarantee equal access for all users.

3 Fundamental Rights Online: Privacy 2.0?

Security technologies have always generated new security risks. Guns can be used to defend against attacks or to enforce compliance with the law, providing trust and security between citizens. However, they can also be used to intimidate a population

[5] On monitoring based on DPI, see, e.g. Bendrath (2009). More generally on monitoring, filtering and blocking, see DeNardis (2007).

into submission and subjection. The monopoly of violence that grounds internal and external sovereignty has been tamed by the legal framework of the Rule of Law that binds those in power to the law, limiting their capability to bend the law to their own purposes. This is a historical artefact that cannot be taken for granted, necessitating a vigilant civil society and an independent judiciary to keep the system of checks and balances in operation. OSTs have a disruptive potential, like guns offline, and require reinstatement of the Rule of Law in cyberspace. In this section, I will discuss the affordances of OSTs with regard to fundamental rights, arguing their radical difference from traditional offline security technologies.

Cybercrime differs from 'ordinary' crime in terms of distance (e.g. remote hacking), scale (e.g. DDOS attack, spam), speed (e.g. dissemination of malware), automation (e.g. Webbots tracing and tracking vulnerabilities, DDOS attack) and interconnectivity (e.g. peer-to-peer file sharing of malicious software, remote hacking; see, e.g. Brenner 2007b). This provides the reasons for treating cybercrime differently from 'ordinary' crime, for instance in terms of lawful investigative techniques. Likewise, OSTs used in the fight against cybercrime differ from offline technique in terms of distance (e.g. remote search; ease of extraterritorial reach), scale (e.g. easy access to data on SNSs, blogs and publicly posted personal information), speed (e.g. deleting evidence on unlawful law enforcement), automation (e.g. profiling of online content and of online behaviours) and interconnectivity (once data are disclosed, difficult to put back in Pandora's box). If both cybercrime and cybersecurity measures have novel impacts on safety and security, the same will probably be true for their impact on fundamental rights. One may expect that the impact of OSTs in terms of distance, scale, speed, automation and interconnectivity change the mode of existence of fundamental rights such as privacy, data protection, non-discrimination, due process and free speech. I will discuss how these novel security technologies reconfigure these rights to prepare for the central argument in the next section, which concerns the image of balance that permeates discussions on liberty and security.

3.1 Privacy in the Era of the Script and the Printing Press

Though some may consider privacy a value, good or interest that is universal, others may claim it is a social construction and determined by cultural norms and values.[6] My own approach is more pragmatic and turns to privacy as a normative practice that is directly related to the infrastructure of information and communication technologies (ICTs) that mediate human practices. I will distinguish four eras defined by their dominant ICT infrastructure: first, the era of the script and that of the printing press; second, that of photography and mass media; third, the database era; and finally, fourth, the era of artificial intelligence and interconnectivity. It would be interesting to start from scratch and discuss the type of privacy that evolves in oral societies, which are face-to-face societies that mainly depend on the spoken word to achieve cohesion—also across subsequent generations (Ong 1982).[7] Though the notion of

[6] On privacy as social construction, see, e.g. Cohen (2012) and Schauer (2001).
[7] See on the introduction of the script and the printing press: Goody and Watt (1963) amd Eisenstein (2005).

critical distance in ethology can be linked to our understanding of privacy as boundary work,[8] it seems obvious that speech as we know it creates a new necessity for such boundary work. Once human animals begin to speak about their inner life, they need to develop normative practices to protect their inner life from invasion or destruction by others.[9] This is crucial precisely to the extent that the inner life of human beings is constituted by language that allows for self-reflection, which may be the precondition for what we call an inner life, though this remains debatable (does my cat lack an inner life, or does it merely lack our kind of language-mediated inner life?). What matters is that our inner life is contingent upon language exchange with others and is indeed constituted by the anticipations of such exchange (a fact coined as double contingency by Parsons and Luhmann and traceable in Ricoeur's discussion of identity construction in *Oneself as another*).[10] The vulnerability generated by this contingency requires us to develop boundaries between self and other, reinforcing a relatively autonomous dynamic of our inner life—notwithstanding the fact that it is constituted and continuously reconfigured by interactions with others.

Privacy as the protection of an inner life constituted by language mediated interaction with others changes its nature with the invention of the script. Ricoeur has wonderfully described how the script triggers a threefold *distantiation* in time and space, based on the *externalisation* inherent in the technologies of the script: a distantiation between author and text, text and reader and author and reader (Ricoeur 1973; Geisler 1985). The audience is enlarged beyond face-to-face communication, the meaning of the text cannot be controlled by the presence of the author, and the text can speak to audiences in other geographical areas and across many centuries. Lévy has performed a similar analysis of the impact of the printing press and the hyperlink, based mainly on various insights from media studies (Lévy 1990). He highlights the notion of the delay as crucial for the era of script and printing press: the distantiations provoked by externalisation (manuscript) coupled with unbridled proliferation (printing press) cause a delay between the act of expressing oneself and that of responding to the expression. This delay disrupts the assumption of a common sense, since it is no longer clear that the audience shares the context of the author. The printing press thus triggers the need for interpretation: the act of deciding the meaning of an expression in differing circumstances. At the same time, the proliferation of different voices that can be accumulated in a library prompts the 'monologue interieur' that some authors trace to the Renaissance period when private libraries became a possibility. Montaigne's essays can be read as the expression of such an interior monologue, which is more of a dialogue in fact. This is where privacy comes in as a kind of side-effect—or rather affordance—of the printing press (Stalder 2002): the retreat from social life into a state of solitary silent reading creates a new dynamic in the inner life of the person. It prepares for a new kind of autonomy, based on the theatre of different voices raised in the back of one's mind by the reading of

[8] On critical distance in animal behaviour, see Hediger (1970); on privacy as boundary negotiation, see Altman (1975).

[9] This may seem a rather strong claim. However, if our inner life is constituted by our capacity to 'speak our mind' after being 'spoken to', then another's knowledge of our inner self may indeed be both invasive and destructive (e.g.Hudson 2005).

[10] On the emergence of an inner life as a result of language mediated interaction, see Wolf (2008), on double contingency, see Parsons (1991) and Luhmann (1996). On identity construction, see Ricoeur (1992).

contradictory arguments. The curious effect of this particular development of the reading brain is that thoughts are developed in the privacy of one's reading mind that are fundamentally opaque to others. These thoughts *can be* but *need not* be expressed in speaking or writing—they can be expressed as private thoughts and though this seems rather obvious to us now, it may have been an emergent affordance of the printing press. A historical artefact. The bottom line of this argument is that it may be the case that privacy in the sense of shielding our developing inner life is a consequence of a particular ICT infrastructure and should not be taken for granted. Novel ICT infrastructures may overrule this kind of privacy, which is closely related to the idea that thoughts are inherently free—even when their expression can be censured.

3.2 Privacy in the Era of Photography and Mass Media

The advent of the ICT infrastructure of mass media threatened our understanding of private life that may indeed have been instituted by the privacy of the bookish mind. At the end of the nineteenth century, Warren and Brandeis wrote their famous article in the *Harvard Law Review* on 'the right to be left alone' in response to the publication of celebrities' photographs that were made and disseminated against their will (Warren and Brandeis 1890). They in fact initiated what is now called the portrait right. This right allows a person to object to her image being published if her private interest against publication outweighs the public interest in publication. Privacy is thus pitted against freedom of information. The 'right to be left alone', which they termed 'the right to privacy', was originally engaged as part of tort law; whoever suffered damage from the violation of her privacy could either sue the perpetrator for compensation or request an injunction, based on a specific tort of trespass. Only later, this right was 'read into' the Bill of Rights as a constitutional right to privacy against government interference. Mass media that can broadcast fragments of a person's private life to a wide audience afford infringements of private life by overruling a person's right to social withdrawal and to her capability to construct and reconfigure boundaries between self and others.

3.3 Privacy in the Era of Databases

In the database era that erupted in the 1950s and 1960s of the last century—before the advent of computerized data servers and interconnected personal computers—the new ICT infrastructure of databases and concomitant information retrieval raised novel privacy concerns. This time the impact on our sense of self derived from the ability of governments to accumulate large sets of data on its citizens, instigating fear of a surveillance society that would jeopardise the freedom of individual persons to reinvent themselves and to choose their own version of a good life. In 1967, Alan Westin published his *Privacy and Freedom* (Westin 1967), firmly rooted in the liberal tradition that abhors government interference with private life and builds on privacy as a fundamental public good that is preconditional for a strong civil society. The advent of a database infrastructure and practices of intercepting and storing private communication leads him to a new articulation of privacy, or a new dimension of privacy as informational privacy: 'Privacy is the claim of individuals, groups, or institutions to determine for ourselves when, how, and to what extent information about them is communicated to others' (Westin 1967, p. 7). Note, first, that Westin goes

beyond individual privacy and, second, that he articulates a claim that comes close to what the German legal doctrine has coined informational self-determination. One could rephrase this claim by stating that privacy is a matter of control over personal data and this easily lends itself to a view of individuals, groups or institutions as sovereigns ruling over data that concern them as if they own them. The control paradigm has many drawbacks. Personal data are non-rivalrous because access to them is not exclusive unless such exclusion is constructed; in fact personal data such as name, address, personal preferences, health, spending capacity or all kinds of observed behavioural data are attributions that only make sense in the interplay between an individual and her environment. The idea that personal data can be owned or can even make any sense outside a relational context is mistaken; just like the idea that the constitution of our inner life could occur outside interaction with others. However, the fact that the self is a product as well as an originator of communication and information exchanges does not imply that privacy or a measure of control of personal information is not important. On the contrary, precisely because the iteration of our identity is contingent upon information and communication exchanges we need to withdraw to reconfigure the boundaries between self, other and world. This also goes for control over information. As long as this is not understood in a fundamentalist framework, control over personal information is preconditional for negotiating the membrane that *separates us from* and *links us to* the environment.

3.4 Privacy in the Era of Smart Environments

By now, the era of databases has expanded into a novel era that combines artificial intelligence with interconnectivity. Databases have become data servers, filled with data collected by artificial agents that crawl the World-Wide Web of hyperlinked information produced by the interconnected computing systems that form the Internet. In their seminal introduction to *Technology and Privacy: Then New Landscape* Agre and Rotenberg define the right to privacy as 'the freedom from unreasonable constraints on the construction of one's identity' (Agre and Rotenberg 2001, p. 7). As explained elsewhere (Hildebrandt 2011), this definition has six advantages over earlier definitions of privacy, covering the physical, decisional, spatial and informational dimensions of privacy. First, it establishes a link between privacy and identity; second, it takes a dynamic view of identity instead of taking it for granted; third, it takes a relational view of privacy and identity instead of emphasizing solipsism; fourth, it highlights the relative nature of privacy by restricting itself to unreasonable constraints; fifth, it thus combines negative freedom (freedom from) with positive freedom (freedom to); and sixth, it acknowledges that privacy and identity do not occur in a vacuum but necessarily emerge from the constraints of a particular context. A more complex ICT infrastructure necessitates a more complex understanding of privacy. If the computational environment tracks and traces our observed machine-readable data across any number of contexts; performs various types of machine learning operations; and then matches the patterns it has mined to our real time behaviours, we need to reconsider how this impacts our inner life and the constitution of the self. It also means that the right to data protection, to non-discrimination, to due process and to free speech become implicated in the right to privacy. They seem to be at stake simultaneously because our smart environments try to pre-empt us by means of refined and mathematically

inferred categorisations: the resulting behavioural sorting allows for invisible discrimination; bypasses the conscious brain and thus precludes opportunities for reflection and contestation that are crucial for due process; and finally it uproots the meaning of free speech since our access to knowledge is filtered by what information the environment 'thinks' we prefer to access (e.g. Google instant). In the context of behavioural advertising McStay speaks of pre-empting our intention, and in the context of increased real time multitasking Wolf speaks of the reconfiguration of the morphology and behaviours of our reading brains (McStay 2011, p. 3; Wolf 2008). The slow construction of a rich inner life based on the sequential process of close reading a multiplicity of books may be overruled by an ICT infrastructure that runs on hyperlinked simultaneity and parallel processing. We may have to reinvent privacy as well as identity.

3.5 Other Implicated Fundamental Rights

The fundamental right to data protection—as codified in the Charter of Fundamental Rights of the EU—is an invention of the database era, necessitated by the advance of automation in database management (http://ec.europa.eu/justice/fundamental-rights/charter/index_en.htm). In its most elaborated form, data protection is about all forms of data processing (including capturing and storing, but also data mining); discriminates between the roles of data subject (whose data are at stake), data controller (who controls the purpose of processing) and data processor (who operates on the data under supervision of the data controller); requires a legitimate ground for processing (e.g. consent, contract and law); stipulates conditions for processing (e.g. purpose specification, data integrity); imposes information obligations for the data controller (e.g. transparency about who is processing what data with what purpose, auditability of its operations); attributes rights to the data subject (e.g. to access, correct or erase her data) and elaborates a right not to be subject to automated decisions if these have a significant impact or legal effect. Liability of the data controller is default and the control paradigm is deeply inscribed in the core principles of purpose binding, consent and data minimisation.

The fundamental right of non-discrimination concerns the prohibition of discrimination in the context of occupation or employment, the provision of goods and services or other important domains of everyday life such as housing, social security or healthcare. Such prohibitions, which vary across jurisdictions, are limited to a set of grounds and do not touch price discrimination based on economic calculation or actuarial approaches to insurance. The grounds also vary across jurisdictions but are similar to those summed up in the EU data protection legislation that prohibits the processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life. The proposed EU General Data Protection Regulation adds criminal convictions or related security measures and genetic data. Exceptions apply, such as explicit informed consent, the vital interest of the data subject or specific legal obligations.[11] Next to non-

[11] Art. 8.5 of the current Data Protection Directive 95/46/EC already sums up a set of conditions that applies when criminal convictions or related security measures are at stake (the same conditions return in art. 9.2(j) of the GDPR. Under the current Directive, however, these data have not been qualified explicitly as personal data whose processing is prohibited.

discrimination law, specific legislation is in force to ensure equal treatment of man and women, often complemented with a justification for positive discrimination meant to address indirect discrimination. This is very important for our subject, because online security measures may generate indirect discrimination on forbidden grounds (Custers et al. 2013; Hildebrandt and Gutwirth 2008). Though, legally speaking, this is not always prohibited, it raises formidable questions about the substance of the right to non-discrimination in the era of computational pattern recognition.

The fundamental right to due process derives from the US Bill of Rights and refers to specific safeguards in the case of governmental intervention that deprive a person of life, liberty, or property. In essence it provides subjects with a right to contest a decision against their interests. Within the European context it can be found in the right to a fair trial. In a more broad sense one can understand due process as the effective right to *be made aware of* and to *be capable of* contesting the violations of other rights. In the context of data mining and profiling is would entail an effective right to be made a ware of automated decision-systems and the envisaged effects they have and the right to object against being submitted to such decision making (Steinbock 2005; Hildebrandt and De Vries 2013). This right has been codified in EU data protection legislation and can be framed as a right to profile transparency: the right to know that one is treated in a certain way on the basis of statistical or individual profiles and the right to contest such treatment. The fact that profiling is usually invisible and hidden behind trade secrets, intellectual property rights or security considerations means that the era of smart environments may deprive this right from its substance. Especially in the context of criminal law or intelligence, we are not enlightened about the computational grounds of being treated one way or another. To the extent that OSTs depend on invisible monitoring, certification mechanisms for encryption or authentication that rely on trust that cannot be verified, or invisible filtering of data streams these OSTs challenge the right to due process in the general and—in case of criminal investigation—the more specific sense.

The fundamental right to free speech can be understood in different ways. It can refer to the duty to refrain from interference, which relates to the freedom from monitoring, filtering and blocking of Internet traffic. This entails a negative obligation imposed on the government. The right to free speech can also refer to the duty to ensure that a private public sphere is sustained that fosters free speech outside the official public sphere or parliament. It is related to the current prohibition to impose obligations of systematic monitoring on ISPs (under EU jurisdiction) and to issues such as Net Neutrality and a proposed universal right to Internet access. EU law prohibits that Member States impose obligations for systematic monitoring on ISPs. The European Court of Justice (EcJ) has already decided that this prohibition rules out that IPSs can be ordered by a court to perform systematic monitoring to filter or block illegal content that has been uploaded in violation of copyright. However, the use of OSTs by ISPs for systematic monitoring of Internet traffic to filter malware and to inspect packages for security threats falls outside the scope of this prohibition if it is a result of their own initiative. In that case the question is whether they are committing a criminal offense by incepting communication—or whether this falls under the exceptions made for technical security measures (Wolk 2012).

3.6 OSTs and Privacy 2.0?

As suggested above, it seems that data protection, non-discrimination, due process and free speech are increasingly implicated in the right to privacy. This does not mean that these rights are equivalent with privacy or merely subdivisions. It does signal that a definition like that of Agre and Rotenberg is more promising as an inroad to the landscape of potential privacy infringements than narrow definitions like those of 'the right to be left alone' or 'the control over one's personal data'. If we need to reinvent both privacy and identity we may as well nourish the awareness of pivotal connections with these other rights—that may require reinvention themselves.

Instead of taking the transformation of the ICT infrastructure and the ensuing need to rethink privacy as a threat, I will take it as a challenge. Privacy has never been a fixed entity and it has managed to survive many types of threats already. But this cannot be taken for granted and requires hard work. Also, the fact that novel types of privacy infringement are at stake does not imply that the earlier understanding of privacy becomes obsolete. The script has not led to an obliteration of speech, the printing press has not outlawed writing, mass media have not eradicated books. We may expect that smart environments will continue to host speech, writings, books and other media. At the same time we must acknowledge that speech was transformed by the dominance of the printed word and, for instance, online social networks have changed the format and function of newspapers and other publications. This suggests that continuity is mingled with discontinuity and the point is to flesh out how this matters. In the context of OSTs we have seen a change in distance, scale, speed, automation and interconnectivity compared to earlier security technologies. Based on the previous analysis of the impact of ICT infrastructures on the substance of privacy, I will briefly explain how OSTs may change the substance of the right to privacy in online environments. Since OSTs are to be situated in the era of artificial intelligence and interconnectivity, they are in fact defined by the transformative quality of this new environment with regard to distance, scale, speed, automation and interconnectivity.

The privacy implication of OSTs can best be defined in terms of the most inclusive definition of the right to privacy, 'the freedom from unreasonable constraints on the building of our identity', that incorporates both the 'right to be left alone' and a measure of 'control over personal information'. *The first type of OSTs*, that aim to ensure confidentiality of communication and rightful online authentication, may infringe privacy in the sense of control over one's personal data when key management or certification go awry and data breaches occur. Attempts to prevent this may involve monitoring that could involve profiling with far more invasive privacy effects. Also, online authentication may – depending on whether this involves full identification or the management of credentials – allow for full scale monitoring behind the access-point. This will, again, allow profiling that has a more complex and more invasive impact on privacy. *The second type of OSTs* aim to detect and counter online threats and vulnerabilities, to filter malicious code or illegal content, or to block IP addresses that engage in attacks, disseminate malicious code or illegal content. All this requires monitoring of Internet traffic and thus links up with surveillance. Monitoring can consist of shallow or deep packet inspection to check on volume, routing, origin or destination, or even content. It affords eavesdropping,

censorship and all kinds of statistical operations such as data mining, for instance for commercial purposes, law enforcement or intelligence. The point is not whether this is actually done or whether current purpose of developing these technologies is to spy, profile or spam people. The point is that all this becomes technologically and economically possible. *The third type of OSTs* aim to fight cybercrime by enabling law enforcement to detect, prosecute, stop and prevent cybercrime. They consist of technologies to gain secret access to computing systems, to capture, observe and/or intercept data and content. These technologies basically enable users to commit various types of crime, and their use by law enforcement requires special competences. Remote hacking, especially extraterritorial law enforcement by means of remote hacking, are high on the agenda of politics. The German Federal Constitutional Court considers remote hacking unconstitutional if not conditioned by a series of safeguards, that should prevent large scale surveillance of the content of computing systems and communication of individual citizens.

OSTs have a significant potential to infringe privacy in the traditional sense of 'the right to be left alone' or to keep 'control over personal information'. The possibility to intercept confidential communications, to gain secret access to computing systems and to engage in invisible systematic monitoring of Internet traffic change the distance from which privacy can be infringed, the scale of the infringements, the speed with which information and communication can be observed and stored, automation renders the implications of data processing opaque because it is not easy to foresee how information can be used against a person and finally, interconnectivity makes privacy vulnerable to unexpected cross-contextual visibilities.

For instance, solitary retirement into the inner life of the reading mind still requires protection against censure of thought, but the nature of the threat is transformed in the age of neuro-marketing (Ariely and Berns 2010; Murphy et al. 2008). We can foresee a time when OSTs profile online behaviours in correlation with knowledge of brain behaviours—to detect criminal intent—and this could have a potentially devastating effect upon privacy. This effect relates to the fact that discrete behavioural data points can be matched against invisible profiles that are kept secret to achieve a new type of 'old school' security by obscurity (Hildebrandt 2011). Like when, in the old days, the charge was hidden from the defendant because disclosing it could provide her with the means to actually defend herself. Similar reasons can be given to keep criminal profiles a secret. Such profiling involves a transformation in distance (matching can be performed remotely), scale (it can mine Big Data of online behaviours), speed (the computational techniques to match individual data points with criminal profiles are incredibly fast), automation (the entire process of 'flagging' potential suspects depends on automation) and interconnectivity (online behaviour reveals behaviour patterns because it crosses contextual borders due to the interconnectivity that defines online social networks and the Internet more generally). The contextual integrity that was default in the era of the printing press with its differentiation between contexts of work, home, church, leisure, politics and economics will in fact necessitate deliberate intervention to survive the current business models that thrive on cross-contextual data mining (Nissenbaum 2010; Cohen 2012). Control over one's personal data does not necessarily solve this problem, because the problem of upcoming OSTs may be that they could enable the correlation of online monitoring output with any kind of profiles sold by large database companies. Such public-private or private-private

collaboration seems highly problematic for the capability of individuals to anticipate how they will be profiled, which is a prerequisite for construction of identity. Other definitions of privacy may have a problem to coin this as a privacy problem, because merely retreating into the privacy of one's mind may no longer work and hiding all the trivial data points that trigger criminal profiles is not possible if you cannot foresee which data match what profiles.

With OSTs, we seem to have developed privacy threats 2.0, and this invites counter infringement measures 2.0. I will briefly return to this point in Section 3.3, arguing for freedom infringement impact assessments for OSTs and the implementation of legal protection by design. First, however, we need to inquire into the nature of the infamous mantra of balancing liberty and security.

4 The Notion of the Scale: Trade-off and Balance?

Immediately after 9/11 the metaphor of the scale gained traction. It gives the impression of balance and reasonableness in times of emergency. The core idea seems to be that between security and liberty, we cannot have our cake and eat it too; choices will have to be made and if so, they had better be well balanced. The idea that some of our liberty must be given up to achieve security has a rhetorical ring that fits the political agenda of extended law enforcement competences, as well as that of the security technologies' industry in the broad sense, including arms production for military purposes. That agenda also includes OSTs and we can imagine that companies developing and selling them have two strategies to enlarge their turnover: they can increase their market share or expand the market. This usage of the metaphor of the scale can be used to sell security technologies which may infringe fundamental rights: if we want online security, we must accept that privacy, data protection, non-discrimination, due process and free speech will have to be limited. Giving up a measure of privacy to gain a measure of security sounds reasonable to many people. In fact, the issue is then framed in terms of a trade-off: the more security, the less privacy. This has been called a zero-sum game.

Another way of looking at the scale is to demand that more security measures that impact civil rights on one side of the scale, require more effective legal safeguards on the other side. This is what balancing is about. In terms of the legal framework that determines the justification for violations of the right to privacy, both metaphors are at stake and I will return to this in the last part of this section under the heading of Section 4.3.

4.1 Three Preliminary Observations

Before embarking on the intricacies of the image of the scale, I will raise three issues. The first concerns what 'stuff' we are balancing, the second suggests that security is not in the same category as fundamental rights and the third inquires into the nature of power and authority as different types of enablers of security measures.

The discourse of balance evokes the notion of weighing. On a scale, one can weigh different ingredients, assuming that all that matters is their weight. In the case of security technologies that infringe fundamental rights we must ask the question what

we are comparing and whether that makes sense. Are we balancing the public good of security against the private good of privacy, or must we acknowledge that both are public goods? And what stuff are these goods made of: rights or interests?[12] Are we then balancing rights to safety or security against rights to privacy or non-discrimination. Or individual rights such as privacy against collective interests such as security? Or private interests in privacy against public interest in security? Obviously such questions are related to traditional distinctions in moral philosophy. Kantian deontologists may vouch for a rights approach, whereby rights are the trump cards that will overrule interests. Benthamite utilitarians may vouch for an approach based on aggregate interests, though Millian liberals could restrict the utilitarian calculation by claiming a fundamental right to liberty that can only be overruled by security interests if they serve to protect individual liberty. I will not move into these discussions but observe that pitting individual rights or interests in fundamental rights against the public good or collective interest in online security is highly problematic. It assumes that fundamental rights and security are at the same level, whereas their relationship is more complicated. For instance, to the extent that one is conditional for the other. It also assumes that privacy is not a public good or a collective interest and security not a private good or individual right. Privacy seems to be, however, both a private interest and a public good and the same goes for security. For their realisation private interests in privacy and security both depend on a legal framework that provides a minimum of legal certainty with regard to these interests, precisely because individual privacy and individual security are public goods. Legal certainty is important because to sustain societal trust and individual flourishing we need more than ethical obligations to respect privacy and provide security. The law has teeth, and this is a necessity to provide for continuity and trustworthiness.

This brings to the fore the question whether legal certainty can be achieved without security in the broad sense of being safe against violent attacks. This relates to the question of whether security is, like privacy, a human right or rather a precondition for the legal framework on which effective human rights depend. Many authors would claim that security in the broad sense is constitutive for the state, and in that sense not at the same level of analysis as fundamental rights which depend on the state to be enforceable (Piret 2008; Hildebrandt 2013). This issue is of great importance for online security, because of the division of tasks between private companies and government agencies. If states delegate responsibility for online security to ISPs one could argue that it remains responsible for the delegation, including the human rights violations it risks. The notion of balance is transformed if security becomes the right on which other rights, notably human rights depend. In that case security technologies that violate privacy and diminish legal certainty with regard to other human rights threaten the constitution of the state itself, because the state fails to achieve the kind of protection on which its authority is based.

This raises the issue of power and authority (Hildebrandt 2010, 2013). These are hefty terms, often used interchangeably though each has its own series of connotations. Authority is based on command and control, power is based on the economic or military ability to enforce submission. Under the Westphalian system of sovereign states that still rules part of international relations and internal sovereignty, state

[12] On different theories concerning the difference between rights and interests (e.g. Edmundson 2004).

authority depends on the monopoly on violence and the power to protect subjects or citizens against violence from other citizens and other states. Social contract theory from Hobbes to Rousseau somehow assumes that the state is capable of providing the fundamental security required to go about one's daily business.[13] Once such security can no longer be provided, the social contract crumbles and states dissolve in civil war. Security in that broad sense, combining safety against attack and legal certainty, seems conditional for state authority. This means that authority works within the state but hardly between competing states and transnational corporations, where power is at stake rather than authority. This raises a number of questions as to how security technologies function across national borders, playing out across various jurisdictions, where the user may be unaware of the actual location of the data servers that are being accessed to wipe out botnets or other threats to cybersecurity. It is worthwhile to note here that the protection of human rights ultimately depends on a court of law and an administration that grants and executes the substance of the right, whereas at the level of extraterritorial jurisdiction such protection is not guaranteed. In that case, security may be provided without being complemented with the safeguards of fundamental rights.

4.2 Six caveats from Jeremy Waldron

Not long after 9/11, when criticizing security measures in the USA was easily interpreted as unpatriotic behaviour, legal philosopher Jeremy Waldron spoke at a workshop on 'Terrorism and the Liberal Conscience' at All Souls College, Oxford on the subject of 'Security and Liberty: The Image of Balance' (Waldron 2003). In his topical paper, he develops six concerns regarding the idea that international terrorism requires states to strike a new balance between security and liberty. These concerns are relevant for similar calls to strike a new balance with regard to online security, based on the plethora of threats and vulnerabilities that endanger critical infrastructure, service providers and individual citizens.

First, the image suggests that by diminishing liberties we will achieve security. This, however, cannot be taken for granted, it will depend on a number of empirical facts, many of which are simply not known when taking the measures. In other words, their effectiveness cannot be assumed but should be assessed. Whether OSTs achieve effective protection against threats and vulnerabilities of the ICT infrastructure and various types of cybercrime does not depend on whether they infringe fundamental rights. In many instances alternative technologies that cause less or even no infringement could be equally or even more effective. *Second*, the image suggests a measure of precision that is inherent in cost–benefit analyses; to increase security (benefit) we must decrease liberty (cost). Such precision is an illusion, due to various types of incertitude, for instance those discussed by Stirling, who discriminates between risk, uncertainty, ambiguity and ignorance. The precision suggested reinforces a bias to quantifiable, machine readable criteria whose weighing depends on a *tertium comparationis* that may not be available: how do we weigh rights'

[13] This is related to the fact that contract theory concerns the institution of the monopoly of violence, that is only acceptable insofar as the state actually manages to protect its subjects from violent attack. On the complexities of, e.g. the right of resistance in Hobbes, Locke and Rousseau (e.g. Johari 1987, p. 388).

infringements against security interests? *Third*, the image suggests that liberties can be traded at will, though from the perspective of the Rule of Law these liberties are exclusionary reasons (Raz), side constraints (Nozick) or trump cards (Dworkin), that cannot simply be overruled by assumed security gains. The image thereby seems to privilege a utilitarian rather than a Kantian ethics. *Fourth*, the image also suggests that one's liberty must be traded against one's security, whereas in practice there is a problem of distribution. Usually the liberties of specific groups (illegal immigrants, unemployed, convicted criminals, suspects, outliers in the case of profiling technologies) are traded against the security of other groups (those assumed to be law abiding citizens, those who remain 'under the radar' in the case of profiling technologies). This raises the issue of distribution of costs and benefits. *Fifth*, the image is not transparent about the fact that by diminishing liberties the powers of the state are increased and that this simple fact will entail security threats. Montequieu's warning about the need to introduce countervailing powers regards precisely this point. OSTs may reduce citizens' security against state powers. *Sixth*, the image may condone symbolic measures that have no real effect for security, hoping to achieve a largely illusionary public sense of security. This has been coined fact free policies and we should note that much of what is presented as evidence-based policies in fact engages a rhetoric that is entirely fact free.[14]

4.3 Limitation of Fundamental Rights

International human rights law as well as fundamental rights granted by national constitutions employ various strategies to limit the scope of fundamental rights, without losing their substance. Limitation is inevitable, either because they clash with public goods that are conditional for the effectiveness of fundamental rights or because various rights or liberties clash and must be aligned one way or another. Within the context of the European Convention of Human Rights which determines the human rights framework within the Council of Europe (52 states), the limitation of the rights to privacy, freedom of thought, conscience and religion, freedom of expression and freedom of assembly and association have a similar decision system to determine whether a security technology measure that violates one of these rights can be legally justified. This decision system can be used to exemplify how the dual meanings of the image of balance can be related to the impact of OSTs on fundamental rights. It has been called the triple test and combines the notion of the trade-off (zero sum) with the notion of the balance (win–win). I will discuss this triple test in more detail in relation to the justification of a limitation (infringement) of privacy by OSTs.

After attributing a right to privacy in the first paragraph of Art. 8 of the Convention, the second paragraph articulates the following three cumulative conditions for a justified, permitted infringement (e.g. Sottiaux 2008; De Hert and Gutwirth 2009): the infringement must be in accordance with the law, necessary in a democratic society and have a legitimate aim. It is important to note that second paragraph of Art. 8 thereby acknowledges the idea of a trade-off. It only applies if the right to privacy is indeed infringed and stipulates that this is only allowed if the infringing

[14] In the context of anti-terrorism measures such as anti-money laundering, see, e.g. Passas (2006).

measure is necessary and proportional to achieve the good of public safety or the protection of public order, health or morals or the protection of the rights and freedom of others. The necessity is understood as a requirement of proportionality between infringing measure and legitimate aim, but ultimately implies that OSTs that cannot be expected to achieve the goal of public safety or one of the other legitimate aims, cannot be justified. As such, a trade-off is a necessary—but not sufficient—condition in the sense that infringement is only allowed if substantial benefit is to be expected that is in proportion to the cost. Also, the necessity is defined in relation to a democratic society and this rules out any measures that aim to achieve public order or public safety in a way that would violate democratic norms and values. In the case law of the European Court of Human Rights (ECHR) necessity is interpreted in terms of a pressing social need that must warrant the measure that is at stake. Again, this pressing social need must stay within the bounds of a democratic society, and thus show equal respect and concern for each member of society; a pressing social 'need' to suppress minorities or engage in prohibited discrimination falls outside the scope of the second paragraph of article 8.

On top of the condition of a proportional trade-off, the second paragraph also requires that the measure is in accordance with the law and this introduces the model of balance. In the case law of the ECHR the requirement of lawfulness is detailed as demanding a basis in national law that is adequately accessible, sufficiently foreseeable and contains effective safeguards. Adequate accessibility and sufficient foreseeability means that citizens can anticipate what types of measures their government can take that might infringe their privacy and thus relates to their reasonable expectation of privacy. For instance, secret surveillance may be allowed, if citizens can foresee in what circumstances it might occur, even if the police will not notify them before wiretapping or requiring access to the personal data of a subscriber to an online service provider. Adequate and effective safeguards necessitate that secret surveillance is neither unlimited in scope (time and content) nor in scale (number of people, frequency of interception) and the condition of safeguards also entails that a warrant or permission from a judge is needed if the invasiveness, frequency or duration of the measure increases beyond a certain threshold. The most important safeguard has been that infringing measures are only allowed in specific cases, ruling out general monitoring of groups or individual citizens.

OSTs that aim to detect and counter online security threats, vulnerabilities and various types of cybercrime may infringe privacy. Building on the triple test of European human rights law, such infringements should be justifiable in terms of a proportional trade-off that instigates a threshold before the employment of such technologies is allowed. If that threshold is reached, the requirement of proportionality demands the implementation of counter infringement technologies that reduce potential infringements to what is reasonable in relation to expected benefits. In that case, the notion of proportionality becomes contingent upon the technical and economic state of the art in counter infringement technologies. Next to that the employment of OSTs should be conditioned by the implementation of a set of safeguards that are proportional in relation to the expected scope of potential infringements (the more substantial the infringement, the more substantial the safeguards). These safeguards will require human intervention, for instance to judge whether the threshold for a justifiable trade-off is reached—while in an automated

environment some of the safeguards should be automated,[15] for instance by flagging potential abuse of OSTs based on certain indicators.

In short, the issue of the trade-off means that if an OST is not effective in reaching a legitimate aim, it cannot be proportionate and must be prohibited in as far as it infringes privacy or other fundamental rights. To assess this, we need freedom infringement impact assessments for OSTs, resulting in evidence-based OSTs instead of a fact free security policy. The issue of balance requires that the more serious the infringement may be, the more substantial the required counter infringement technologies should be. This can be achieved, for instance, by developing smart security and data protection by design—as imposed by the proposed General Data Protection Regulation.

5 Concluding Remarks

The reach, scope, speed, automation and interconnectivity of online threats, vulnerabilities and various types of cybercrime have offset a growing demand for OSTs. Various types of OSTs can be distinguished, based on encryption technologies and key management, on monitoring, filtering and blocking and on hacking and remote hacking of computing systems. OSTs are meant to increase online security in the broad sense of safety from attacks against confidentiality, integrity and availability of personal data and critical infrastructure; resilience against fraud and forgery; and detection and prevention of child pornography and violations of copyright. At the same time, they afford violations of human rights, notably privacy. In this paper, I have inquired into the nature of privacy in the era of artificial intelligence and interconnectivity and estimated the potential impact of OSTs on our privacy, suggesting that a more generic understanding of privacy and identity is required than either 'the right to be left alone' or 'control over personal information'. Agre and Rotenberg's definition of the right to privacy as 'the freedom from unreasonable constraints on the construction of one's identity' does a better job in pinpointing what is at stake after the transformation of security technologies that function in the online world. Their reach, scope, speed, automation and interconnectivity turn OSTs into a powerful and largely invisible threat to a reasonably independent development of our identity and this threat should inform decisions on the investment in and employment of OSTs.

In this contribution, I inquire into the nature of the metaphor of the scale, fleshing out the difference between a trade-off and a balance. A trade-off would mean that more security means less privacy; a balance would mean that the more privacy invasive an OST the higher the threshold should be for allowing it, the more counter-infringement technologies are required, and the more effective legal safeguards must be implemented. As to the trade-off I discuss a number of issues around the idea of a trade-off, for instance that of distribution: the privacy of some people is traded against the security of others. Also, we must remember that security technologies that are not effective cannot be a justification of a privacy infringement, even

[15] This involves Legal Protection by Design and/or Privacy by Design (see, e.g. Hildebrandt 2011; Langheinrich 2001).

when thinking in terms of a trade-off. The notions of a trade-off and a balance have been integrated in the decision system of the right to privacy in the European Convention of Human Rights that requires a triple test for security measures that infringe privacy. This test provides an interesting framework to think about balancing OSTs with fundamental rights; it has the advantage of consolidating decennia of case law in which weighing the demands of safety and security against the requirements of fundamental rights within the framework of constitutional democracy.

References

Agre, P. E., & Rotenberg, M. (2001). *Technology and privacy: the new landscape* (3rd ed.). Cambridge: MIT.

Altman, I. (1975). *The environment and social behavior. Privacy personal space territory crowding*. Monterey: Brooks/Cole.

Ariely, D., & Berns, G. S. (2010). Neuromarketing: the hope and hype of neuroimaging in business. *Nature Reviews Neuroscience, 11*(4), 284–292. doi:10.1038/nrn2795.

Bendrath, R. (2009). Global technology trends and national regulation. Explaining Variation in the Governance of Deep Packet Inspection. Conference paper at International Studies Annual Convention New York City, 15–18 February 2009. http://userpage.fu-berlin.de/bendrath/Paper_Ralf-Bendrath_DPI_v1-5.pdf

Bennett, C. J., & Lyon, D. (2008). *Playing the identity card: surveillance, security and identification in global perspective*. London: Routledge.

Brenner, S. W. (2007a). History of computer crime. In K. De Leeuw & J. Bergstra (Eds.), *The History of Information Security* (pp. 705–721). Amsterdam: Elsevier

Brenner, S. W. (2007b). *Law in an era of 'smart' technology*. New York: Oxford University Press.

Cohen, J. E. (2012). *Configuring the networked self: law, code, and the play of everyday practice*. New Haven: Yale University Press.

Custers, B., Calders, T., Schermer, B., & Zarsky, T. (Eds.). (2013). *Discrimination and privacy in the information society*. Berlin: Springer.

De Hert, P., & Gutwirth, S. (2009). Data protection in the case law of strasbourg and luxembourg: constitutionalism in action. In S. Gutwirth, Y. Poullet, P. De Hert, S. Nouwt, & C. de Terwangne (Eds.), *Reinventing data protection?* (pp. 3–44). Dordrecht: Springer.

DeNardis, L. (2007). A history of interent security. In K. De Leeuw, J. Bergstra (Eds.), *The History of Information Security* (pp. 681–704). Amsterdam: Elsevier.

Edmundson, W. A. (2004). *An introduction to rights*. Cambridge: Cambridge University Press.

Eisenstein, E. (2005). *The printing revolution in Early Modern Europe*. Cambridge: Cambridge University Press.

Gandy, O. H., Jr. (2000). Exploring identity and identification in cyberspace. *Notre Dame Journal of Law, Ethics & Public Policy, 14*, 1085.

Geisler, D. M. (1985). Modern Interpretation theory and competitive forensics: understanding hermeneutic text. *The National Forensic Journal, III*, 71–79.

Goody, J., & Watt, I. (1963). The consequences of literacy. *Comparative Studies in Society and History, 5*(3), 304–345.

Gürses, S., Berendt, B., Santen, T. (2006). Multilateral security requirements analysis for preserving privacy in ubiquitous environments. In *Proceedings of the UKDU Workshop* (pp. 51–64).

Hediger, H. (1970). *Man and animal in the zoo: zoo biology*. London: Routledge.

Hildebrandt, M. (2010). The indeterminacy of an emergency: challenges to criminal jurisdiction in constitutional democracy. *Criminal Law and Philosphy, 4*(2), 161–181.

Hildebrandt, M. (2011). Legal protection by design. *Legisprudence, 5*(2), 223–248.

Hildebrandt, M. (2013). Extraterritorial jurisdiction to enforce in cyberspace. *Toronto Law Journal, 63*(2), 196–224 (Focus Feature: Criminal Jurisdiction: Comparison, History, Theory).

Hildebrandt, M., & De Vries, K. (Eds.). (2013). *Privacy, due process and the computational turn: the philosophy of law meets the philosophy of technology* (pp. 196–224). Abingdon: Routledge.

Hildebrandt, M., & Gutwirth, S. (Eds.). (2008). *Profiling the European citizen. Cross-disciplinary perspectives*. Dordrecht: Springer.

Hudson, B. (2005). Secrets of the self: punishment and the right to privacy. In E. Claes & A. Duff (Eds.), *Privacy and the criminal law* (pp. 137–161). Antwerp: Intersentia.
Johari, J. C. (1987). *Contemporary political theory: new dimensions, basic concepts and major trends*. New York: Sterling.
Langheinrich, M. (2001). Privacy by design—principles of privacy-aware ubiquitous systems. In *Proc. 3rd. Int'l Conf. Ubquitous Computing* (pp. 273–291). Berlin: Springer.
Leawoods, H. (2000). Gustav Radbruch: an extraordinary legal philosopher. *Journal of Law and Policy, 2*, 489–516.
Leeuw, K. M. M. de, Bergstra, J. (Eds.) (2007). *The history of information security: a comprehensive handbook*. Amsterdam: Elsevier.
Lévy, P. (1990). *Les technologies de l'intelligence. L'avenir de la pensée à l'ère informatique*. Paris: La Découverte.
Luhmann, N. (1996). *Social systems*. Palo Alto: Stanford University Press.
McStay, A. (2011). *The mood of information: a critique of online behavioural advertising*. New York: Continuum.
Murphy, E. R., Illes, J., & Reiner, P. B. (2008). Neuroethics of neuromarketing. *Journal of Consumer Behaviour, 7*(4–5), 293–302.
Nissenbaum, H. F. (2010). *Privacy in context: technology, policy, and the integrity of social life*. Stanford: Stanford Law Books.
Ong, W. (1982). *Orality and literacy: the technologizing of the word*. London/New York: Methuen.
Parsons, T. (1991). *The social system* (2nd ed). London: Routledge.
Passas, N. (2006). Fighting terror with error: the counter-productive regulation of informal value transfers. *Crime, Law and Social Change, 45*(4–5), 315–336.
Piret, J.-M. (2008). Politics, sovereignty and cosmopolitanism in times of globalisation. *Archiv Fur Rechts-Und Sozialphilosophie, 94*(4), 477–497.
Radbruch, G. (1950) Legal Philosophy. In: E. Lask and C. Wilk (eds.), *The legal philosophies of Lask, Radbruch and Dabin*. Translated by Kurt Wilk, with an introduction by E. W. Patterson (pp. 43–224). Cambridge: Harvard University Press
Ricoeur, P. (1973). The model of the text: meaningful action considered as a text. *New Literary History, 5*(1), 91–117.
Ricoeur, P. (1992). *Oneself as another*. Chicago: The University of Chicago Press.
Schauer, F. (2001). Free speech and the social construction of privacy. *Social Research, 68*(1), 221–232.
Schneier, B. (2006). *Beyond fear. Thinking sensibly about security in an uncertain world*. Berlin: Springer.
Sommer, P., Brown, I. (2011). *Reducing systemic cybersecurity risk*. (IFP/WKP/FGS(2011)3) OECD.
Sottiaux, S. (2008). *Terrorism and the limitation of rights. The ECHR and the US Constitution*. Oxford: Oxford University Press.
Stalder, F. (2000). Beyond constructivism: towards a realistic realism. A review of Bruno Latour's Pandora's Hope. *The Information Society, 16*(3), 245–247.
Stalder, F. (2002) The failure of privacy enhancing technologies (PETs) and the voiding of privacy. In *Sociological Research Online*. 7 (2).
Steinbock, D. J. (2005). Data matching, data mining and due process. *Georgia Law Review, 40*(1), 1–84.
Symantec, Internet Security Threat Report (2011) Trends (volume 17, April 2012). Available at http://www.symantec.com/threatreport/. Accessed 9th November 2012
Waldron, J. (2003). Security and liberty: the image of balance. *Journal of Political Philosophy, 11*(2), 191–210. doi:10.1111/1467-9760.00174.
Warren, S., Brandeis, L. D. (1890). The Right to Privacy. In *Harvard Law Review*. 193(7).
Watney, M. (2012). Cybercrime regulation at a cross-road: state and transnational laws versus global laws. In: *2012 International Conference on Information Society (i-Society)* (71–75).
Westin, A. (1967). *Privacy and freedom*. New York: Atheneum.
Wolf, M. (2008). *Proust and the squid: the story and science of the reading brain*. London: Icon Books.
Wolk, S. (2012). *Filtering and blocking of copyright infringement works—a European perspective*. Rochester: Social Science Research Network (SSRN Scholarly Paper Nr. ID 2186751).
Zedner, L. (2009). *Security*. London: Routledge.