Surveillance as public matter

Revisiting sousveillance through devices and leaks

van der Velden, L.C.

Publication date
2018

Document Version
Other version

License
Other

6 Leaky apps and data shots

Technologies of leakage and insertion in NSA surveillance

6.1 Introduction: Surveillance devices in the public domain

This sixth chapter is about surveillance theory after the NSA disclosures in 2013.64 In the thesis, device-centred approaches have been helpful to explicate the translations and performative effects of devices that make surveillance public. The last case study is set up a bit differently, because it does not focus on one particular project or method of making public. The NSA disclosures stand in a long tradition of whistle blowing, and in her documentary about Edward Snowden, 'Citizen Four', Laura Poitras (2014) shows that in order to make highly classified information public in the digital age, one needs to go through various instances of secrecy: one needs to use encryption, air-gapped machines (computers with no internet connection) and identity verification in hidden channels (Citizen Four). Such social-material trajectories of leaking would be a worthy topic of research. However, with this text I aimed to make a contribution to surveillance theory after the NSA disclosures and that's why it focuses mainly on their content. But like the previous chapter, it tries to produce insights from surveillance material that is made public. The NSA disclosures have brought specific objects, surveillance devices, into circulation, and this case study offers a reflection upon the consequences of that event.

Others have written about the making public of 'secrets' or classified knowledge (Galison 2005) or about how the recent disclosure of NSA's surveillance devices can be understood from a historically informed perspective on surveillance and/or censorship (Galison 2016; Murakami Wood and Wrie 2015). What I have done is taking the event as a possibility to construct a collection of surveillance devices, in order to rethink surveillance theory 'post-snowden', as

64 A previous version of this chapter has been published as "Leaky Apps and Data Shots: Leakage and Insertion in NSA-surveillance" in *Surveillance & Society* (2015).

Surveillance Studies scholars called for after the disclosures (Murakami Wood and Wright 2015). The study is an attempt to theorise a set of leaked data: that of the NSA files and more particular methods for data collection they refer to. 'Collect it All' has become one of the key terms to characterise the practice of the NSA since Edward Snowden's revelations (Greenwald 2014). The apparent totality of surveillance has influenced the way we conceive of surveillance. With this case study I delve deeper into the specific methods of capturing data that have been disclosed. Its central focus is on displaying and analysing a range of surveillance technologies that came to the fore through the NSA files and have thus become present in the public domain.

The disclosures raised various socio-technical, political and economic issues. To name but a few: the lack of oversight, the enormous budgets involved, bilateral agreements between intelligence agencies, the secret decryption lobby by intelligence agents, and the potential information overload for the NSA (Greenwald 2014; Landau 2013; Poitras et al. 2013). The reason I focus on the tools is because the NSA files offer an occasion to reflect upon how we think and talk about technologies in surveillance theory. The disclosures allow me to return to the trajectory of thought within surveillance studies that has revised notions of surveillance after Foucault (see §1.4). Therefore the files provide opportunities to shift our vocabulary about the vague term 'data collection' towards a more accurate conceptual repertoire. In short, this is a good occasion to test whether the conceptual apparatus that is used to make sense of surveillance is capable of accommodating these disclosures about the state of the art of (state) surveillance.

In the next paragraph I will briefly sketch how information and communication technologies have affected concepts of surveillance in the past and how the NSA files have led to discussions, between academics and in public forums, about how contemporary surveillance society should be understood. By drawing on a set of news publications about the NSA files, I will highlight two ways in which the NSA, or occasionally the GCHQ, has captured data. The first consists of a list of programs that extract data due to the fact that internet devices 'leak' data. The second is a list of 'devices that insert', for instance, malware. This is done in order to conceptualise two distinct forms of movement, leakage and insertion, by which data are captured by NSA programs. I will discuss the (theoretical) questions for each form that these devices could raise by drawing on existing media theory and conclude by pointing out future sites of research for Surveillance Studies.

6.2 Information technologies and surveillance theory

Since the rise of popular internet use, commentators have conceptualised the surveillance potential of computer networks. As also highlighted in the first chapter, part of that effort consisted of debunking old notions of surveillance that are not able to deal with contemporary information technologies. In the academic sphere, we have seen consistent criticisms of famous concepts such as Orwell's 'Big Brother' and the Foucauldian panoptic diagram. According to influential scholars in the mid-2000s, these were outmoded ideas because they primarily focused on individual bodies being subjected to techniques of direct supervision. They were considered to be inadequate to grasp the more volatile and fluid movements of data that networked information technologies allow (Deleuze 1992; Haggerty and Ericson 2000; Haggerty 2006; Lyon 2006). Gary Marx's table of the 'new surveillance' (Marx 2002) is indicative of an emerging shift towards the idea that surveillance consists of a complex and remote digitally mediated network that operates far beyond individual bodies and the visual.

Within Surveillance Studies, the work of Foucault especially has been a source of inspiration and criticism. This reworking of Foucault can be traced back to Deleuze (1992) who outlined how disciplinary techniques had been superseded by (code-based) technologies of control. A decade later, Haggerty argued that many surveillance scholars kept imposing panoptic tropes into contemporary reality (Haggerty 2006) (see also §1.4). Presently, a post-Foucauldian paradigm within Surveillance Studies seems to have stabilised itself.[65] Instead of scrutinising the techniques by which physical bodies are being watched, many have shifted attention to the topic of 'data bodies' or online profiles: data composites that relate to our technologically mediated behaviour and that become interesting material for both state and corporate actors (Critical Art Ensemble 1998). Data bodies have shaping effects; they can precede who we become ahead of our actions (Stalder 2002). Terms such as 'surveillant assemblage', pushed by Haggerty and Ericson (Haggerty and Ericson 2000) who took inspiration from Deleuze's concept of distributed agency, describe how persons are disassembled and recomposed into data doubles through fluid network flows. (See also chapter one, §1.4.)

The move towards the notion of the assemblage exemplifies how new concepts also change the sites of research and the questions that are being asked. As highlighted in §1.4, moving from the panopticon to the assemblage means that the locus of power shifts from the disciplinary

65 For critical perspectives on the way this is embraced, see Murakami Wood (2007) and Caluya (2010).

institutions to what Latour has termed 'centers of calculation' (Haggerty and Ericson 2000, 613). These are the places of intersection in which data are reassembled and made sense of, such as digital databases. Hence, it is the database and algorithmic power that becomes of vital interest Scholars have stressed concerns about algorithmic power and its potential effects such as social sorting, in which online profiling can have discriminatory consequences (Lyon 2003), as some profiles are more 'sensitive' than others (King and Jessen 2010, 608).

As also mentioned in the §1.4, the merits of the concept of the assemblage have been questioned. Critics have argued that the very definition of the surveillant assemblage as an agent itself leads away from an understanding of the specificity and agency of individual surveillance practices (Prainsack and Toom 2010, 1119). Haggerty himself expressed some hesitancy about 'successor models' in surveillance theory being stretched to all surveillance contexts (2006, 39). Accompanying his critique he suggested that scholars should take up the questions that 'technologies raise themselves' (Haggerty 2006, 32). One could argue that this becomes difficult when a notion of distributed agency becomes the general guidance to a wide variety of surveillance practices.

An important task, rather than doing away with conceptual models altogether, is to keep in mind how the concepts we use determine the range of questions we ask and do not ask. For instance, shifting attention to data flows and algorithmic calculation risks leaving the specific tools, which capture, data conceptually unattended. Despite the importance of theorising the way power is exerted through data bodies, one risks seeing data as something that can be 'picked up' from data flows in some undefined way, as if people just 'leave' all these pieces of information out there as traces in public space. As other authors have argued, what falls under the term 'data' is too often taken for granted (Gitelman and Jackson 2013), as well as what constitutes a 'data trace' (Reigeluth 2014). Analogously, the devices involved in producing these data traces risk receiving relatively little attention in the process of 'concept production' (Lovink and Rossiter 2012) when the concepts are already in place.

One of the characteristics of the 'new surveillance' according to Gary Marx (2004) was the relative invisibility of the surveying actors and their methods. Surveillance Studies scholars have seen it as their task to unveil the very technologies and objects that make up surveillance networks/assemblages and there are plenty of case studies to be found that pursue this goal (Adey 2004; Ball, Haggerty and Lyon 2012; Elmer 2004; Martin, van Brakel and Bernhard 2009; Monahan 2006). However, technologies of data collection often play a case-specific role in what is subsequently conceptualised as larger 'networks',

'assemblages', or 'sorting machines'. I take a different route and explore whether specific technologies, a variety of which are archived in the NSA files, require their own concepts as well.

6.3 NSA files as game changers?

The revelations about NSA surveillance have obviously changed the landscape of Surveillance Studies. In "After Snowden: Rethinking the Impact of Surveillance", Bauman and his co-authors argue for a rethinking of the 'canons' of Surveillance Studies and critical security studies:

> Scholars have promoted promising ways of thinking about the complex and rhizomatic character of interconnected networks of surveillance tools, in which self-exposure has become commonplace, but now also have to make sense of the extension of intelligence gathering together with powerful integrative platforms such as PRISM. (Bauman et al. 2014, 124)

In other words, they say that the NSA files introduce something new to a specific strand of surveillance theory. Intelligence practices (that might have been very familiar to a particular discipline of intelligence studies) have now become highly visible for those working on everyday surveillance society. The question becomes: In what way are concepts of rhizomatic, and less hierarchal, networks (Haggerty and Ericson 2000) and notions of self-exposure and 'participative surveillance' (Albrechtslund 2008) sufficiently able to make sense of this pile of technologies of intelligence?

Next to academic debate, questions about what the NSA files mean for understanding surveillance have also been posited in public discourse. The NSA affair has sparked a discussion about the appropriate concepts to describe surveillance. For instance, the PEN American Center (2014) did a study into popular metaphors used to describe surveillance (in over 60 news outlets).[66] The Dutch online news magazine *De Correspondent* composed a series of discussion posts about surveillance metaphors in books and films (Berger 2014). My impression from those publications is first that in popular imagery the concepts used to describe what surveillance is 'like' heavily depend on grand systematic and visual concepts such as Big Brother and the Panopticon.[67] Second, the PEN-study highlights that when speaking of what surveillance 'does', words are used which express large-scale

66 PEN is an organization promoting literature and freedom of expression.

67 See also for instance the alternative proposal by Rosen and Santesso to speak of 'The Eye of Sauron' in: *Slate* on 17 July 2013.

movements, such as 'sweeping', but also many shipping and fishing-metaphors such as 'trawling'. In short, the bigger picture is sketched in imaginative ways. The imagination is, however, considerably less visible when the tools employed by these 'Big Brothers' and 'Panopticons' are described. There are surprisingly few metaphors that are used to explain how tools work. How to imagine those devices? They seem to be just too complex, too technical, and too varied.

This writing is an attempt to merge the two issues mentioned above, to draw theoretical consequences from the presence of intelligence technologies in the public domain, and sketch a different imagination by investigating the characteristics of the devices with which the NSA documents present us. I will list methods of capturing data and present ways of looking at the technologies involved. On the basis of this overview, I suggest that we may reconsider the sense of 'remoteness' that is present both in the idea of the 'new surveillance' (Marx 2002) as well as in vocabularies on data flows and algorithmic politics. Although contemporary surveillance technologies do operate from distance, the way in which a number of devices come close to people or networks challenges us to rethink proximity and intimacy. The notions of leakage and insertion serve as illustration of how we might approach these issues.

6.4 Devices for capturing data: Leakage and insertion

6.4.1 Concepts for capturing data

In this study I make use of two concepts for mapping ways by which data are captured: leakage and insertion. I find them useful for several reasons. When looking at the reporting about the NSA it struck me that technologies for data collection were framed within a particular language. Words such as 'access' 'vulnerabilities' 'injections' came up, and also things such as 'unencrypted' 'updates' and 'streams'. They all refer to very material aspects of networking. Especially the first class of words is familiar to people working on security, but to me they seemed more alien to how Surveillance Studies literature writes about surveillant assemblages.[68] Having Latour's suggestions of infra-language (Latour 2005b, 30) mind (see also §2.1.1), I came up with these two concepts aiming to express two dimensions of data collection, in a way

68 A more fundamental reason might be that Surveillance Studies has been inspired by Foucauldian theory that emerged from looking at practices in prisons and medical institutions, and which has not focused on practices of intelligence (e.g. wire tapping, infiltration).

that remains close enough to how these devices were originally talked about in the documents themselves. Both have a connotation that calls into mind the material aspects of apparatuses. Furthermore, the bodily, almost sexual, overtones make us aware of the intimate relationship that surveillance technologies produce—see for instance Wendy Chun (2013) on 'leaky and promiscuous devices'.[69] Whereas 'leaking' has a connotation with a human intention of leaking data, I propose to use the word 'leakage' because it places the activity of leaking in the material apparatus.[70] Similarly, as I will argue, 'insertion' aims to discuss the proximity of surveillance practice as (malicious) software closely interacts with the machines people use. In theorising 'insertion', I draw upon earlier interpretations of viruses and malware by Jussi Parikka (2007). Inspired by Chun and Parikka, I will approach a series of recent controversial surveillance technologies with these concepts.

Leakage' refers to the basic way in which the NSA is fed by bulks of data that are already circulating; data can be picked up by tapping into the flow. Besides the highly visible social media platforms that provide intelligence agencies with easy access to feeds and public profiles, there is a plenitude of unencrypted internet traffic that does not seem to be public but which can be easily accessed. An interesting example is a widespread use of what has been dubbed 'leaky apps': apps that share personal data. According to a collaborative piece by *Propublica*, *The New York Times* and *The Guardian*, the NSA and GCHQ have been trading methods for 'vacuuming up address books, buddy lists, phone logs and the geographic data embedded in photos when someone sends a post to the mobile versions of Facebook, Flickr, LinkedIn, Twitter and other services' (Larson, Glanz and Lehren 2014). As explained in more detail in *Propublica* (ibid.) and *The Washington Post* (Soltani and Gellman 2013a), this information is subsequently used by the NSA and GCHQ in systems for algorithmic sorting. This kind of data is often seen as a 'by-product' of transactions (Savage and Burrows 2007). However, seeing apps or practices as possessing a degree or amount of 'leakage' instead of producing a by-product gives us a different way of conceptualising unwanted data flows. As I will discuss below, leakage might be more fundamental to their workings than by-products.

69 Wendy Chun has pointed to the sexualized vocabulary of talking about network devices in computer science, expressing the fact that these devices need to get in contact with each other to communicate. She describes these indiscretions as 'promiscuous'. For other approaches to leaky technologies see Lyon's notion of 'leaky containers' (Lyon 2001) and Rogers' work (2009a) on leaky 'walled gardens'.

70 See for instance leakage in chemistry, electronics or in audio, in which material or signals are lost or picked up by different devices (Wikipedia 2015).

'Insertion' refers to the way in which the NSA targets people or groups by sophisticated interception methods. Take for instance what has been termed 'Deep Packet Injection' (versus 'Deep Packet Inspection' or DPI). Deep Packet Injection is a method of inserting data into internet traffic when somebody launches a website. As explained in *Der Spiegel*, the NSA uses software to fake and insert a website, thereby diverging the URL-request to a server that is owned by the NSA (FoxAcid), located at key internet switching points (SPIEGEL Staff 2013a). This only works when the server of the NSA is faster than the server that hosts the website that was originally requested; therefore these machines are also referred to as 'Race Servers'. If the intervention succeeds and the person loads data from the fake server, malware can be installed on his or her computer.[71] So whereas Deep Packet *Inspection* is a technique to monitor and filter network traffic, Deep Packet *Injection* actively inserts data in order to capture information or install malicious software. *Der Spiegel* reported that the method is referred to as 'QuantumInsert' and is associated with a particular NSA program, TURBINE, conducted by an NSA division called ANT (Appelbaum, Horchert and Stöcker 2013). This form of attack was used by GCHQ to target the networks of Belgacom (through fake LinkedIn accounts) (SPIEGEL Staff 2013b). Another example of technologies of insertion is covert implants. One document, published in a Dutch newspaper, *NRC Handelsblad*, showed that the NSA intruded 50,000 machines/networks through covert implants (Boon, Deriz, and Modderkolk 2014), pieces of malware able to infect computer networks.

These examples illustrate two different ways, or movements, by which agencies such as the NSA and GCHQ can acquire data: on one end of the spectrum there is the kind of 'leaky technologies' which people use on a massive scale, that as a consequence produce data which can be tapped. The example of 'leaky apps' seems to fit well with conceptual approaches that emphasise the 'remoteness' of contemporary mass-surveillance and the participative dimension by which people massively share personal data. However, perhaps equally important, are technologies that come close to the machines that people use to acquire information that is not in the open. They conduct, arguably, 'on the scene' surveillance (albeit done by non-humans) and complicate the participatory dimension and, in a way, the remoteness of surveillance as well. They interfere with or are literally 'planted in', people's networks and/or devices.

71 The Intercept (2014) gives a visual explanation in: "How the NSA Secretly Masqueraded as Facebook to Hack Computers for Surveillance."

6.4.2 Listing devices

The next few paragraphs try to order devices mentioned in the files along these two lines of movement. The idea of listing aspects of surveillance was originally inspired by Gary Marx's table of the 'new surveillance' who listed shifts in 'dimensions' of surveillance (Marx 2002). In what follows I will be more modest by simply starting with a list of devices that we know are part of recent developments in surveillance technology. This study covers the first year of the Snowden revelations. At the time of the study, there was no public archive of the Snowden Files ("Snowden Files For All"), although people have recently been working on building them (see the "Snowden Surveillance Archive"). The lists of devices I composed consist of a number of tools mentioned in the 'Timeline of NSA Domestic Spying', which was composed by the Electronic Frontier Foundation (EFF). As the EFF collects coverage from a huge number of news outlets, this is a very useful repository. The EFF, being dedicated to 'defending civil liberties in the digital world' (EFF 2017) can be considered an 'issue-expert' in the field of surveillance. The advantage of this 'expert authored list' (Rogers 2013, 48; see also §2.4.3) versus a query in a database such as LexisNexis, which is the standard news database for academic research, is the fact that the EFF's repository provided an option to demarcate news on the basis of 'leaked documents' (which contained over 70 news outlets), a selection criterium that LexisNexis does not offer. That means that within the coverage of the NSA affair, this selection deals with items that focus on the content of the files (versus, for instance, Snowden himself).

I manually coded the tools and programs by type, such as 'access request', 'malware', 'court ruling', 'storage program', 'data analysis program' and then zoomed into those items that explicitly addressed the issue of *data collection*. What I therefore have not highlighted are a range of systems for data analysis and data storage systems. Also excluded from the analysis are methods for data acquisition that fall within the range of (secret) court orders and intelligence exchanges.[72] I selected examples of technologies of leakage and examples of programs that actively intrude networks or computers. The idea was to see whether, when assembled, they raise particular questions. Listing devices can lead to new questions and concerns because, as argued earlier in this chapter, surveillance studies have powerful concepts (the panoptic, the assemblage) in which the devices have a place. This exercise is an attempt to turn that around, and see whether starting conceptualising from the workings of a set devices helps crafting a different outlook in terms of research questions (and possibly research sites).

72 Systems for data analysis are the Internet Monitoring Engine Xkeyscore for querying e-mails, browser history, IP addresses, and keywords. Data storage systems include SOMALGET for audio data.

Obviously, as I have based this on a dataset of news articles, this approach does not deliver exhaustive coverage, nor would the original files be exhaustive. It is a selection taken from a particular context, in a particular time period, and translated by journalists who make choices about what to disclose. According to Laura Poitras and co-authors, the Snowden revelations provide a 'digital snapshot of the world's most powerful intelligence agency's work over a period of around a decade' (Poitras et al. 2013). Whether this repository offers proof concerning cases of practical use is not part of my analysis. I approach this 'digital snapshot' as an example of how surveillance technologies have been 'made public' (Latour and Weibel 2005), thereby fully acknowledging that this has happened with a specific vocabulary, through specific methods and with particular aims.

6.4.3 Insertion

The tools that I clustered under the header 'technologies of insertion' emerge in many different kinds. Table 3 (next page) lists a few examples, including things such as the insertion of vulnerabilities or back doors ('Dual Elliptic Curve') and data shots, or Deep Packet Injection (as explained earlier in the chapter). The middle column shows the phenomenon (for instance 'embed backdoor'; 'bug in fax machine'), the right column mentions the associated name of the program (sometimes there are multiple), and the left column summarises the techniques (for instance, 'inserting vulnerabilities'; 'the automation of insertion').

Table 3. Examples of insertion methods in the NSA files.
The middle column contains the phenomenon, the right column the associated name of the program, and the left column a summary of techniques.

Insertion		Term of reference or (code) name of the tool or program
Inserting vulnerabilities	Singling out encryption keys/certificates vulnerable to cracking by supercomputers Embed flawed formula / backdoor in RSA encryption software	Cheesy Name; Bullrun Dual Elliptic Curve
Placing bug at endpoint devices	Bug in fax machine (implant cryptofax)	Dropmire
Data injection	Giving data 'shots' (packet injection attack) Man in the middle attacks Masquerading websites such Facebook and LinkedIn	QuantumInsert/TURBINE "Flying Pig" and "Hush Puppy (CCHQ) QUANTUMHAND; Operation Socialist
Methods of planting malware	Malware servers Through browser vulnerabilities Through hardware interception (shipping deliveries) Using system admins as access points for malware	FOXACID Interdiction; TAO workstations I hunt sysadmins
The automation of insertion	Scale up implants	TURBINE
The repurposing of web objects	cookie hijacking (using a Google PREF-cookie to target user) Use Plugins to take over computer	 UNITEDRAKE
The platform-dimension of malware	Building plug-ins on top of implants	UNITEDRAKE; CAPTIVATEDAUDIENCE; GUMFISH; FOGGYBOTTOM; GROK; SALVAGERABBIT

From this list, a few issues stand out. One of the more striking things is that computational processes do not only play a role in the 'analysis' of data, or the 'algorithmic' side of knowledge production, as we know very well from Surveillance Studies, but also in the methods of reaching targets. In other words, 'close observation' by non-human infiltration is automated and scalable. For example, when reporting on TURBINE in *The Intercept* on 12 March 2014, Gallagher and Greenwald wrote that the network of implants managed by the NSA's Tailored Access Operations (TAO) had increased from 100-150 implants in 2004 to tens of thousands in a period of 6-8 years (Gallagher and Greenwald 2014a). The documents suggest that TURBINE enables 'industrial-scale exploitation' (citation by *The Intercept* (ibid.)). They explain TURBINE as being 'designed to "allow the current implant network to scale to large size (millions of implants) by creating a system that does automated control implants by groups instead of individually"' (ibid.).

Instead of seeing this, for instance, as a dragnet, we are in need of a term that can conceptualise large-scale, penetration of and insertion into targeted networks. Whereas many scholars in Surveillance Studies have highlighted mass data collection, big data and the automated analysis of big datasets (profiling, discriminatory algorithms and so on), the NSA documents point just as well to large-scale, automated practices of insertion of a wide variety of implants and malware, the effects of which call for further study and conceptualisation. Insertion is a way to describe this other genre of surveillance practices. Furthermore, as I will show later, tropes of insertion are part of a different, more intimate, way of imagining the effects of surveillance.[73]

The above-listed practices of insertion might also pose new questions about the targets of 'sorting'. With this I mean not only the selection criteria that are used to decide whether people under surveillance are 'suspect' or not, but also the selection process in which it is decided whether or not someone is useful for the insertion of malware. For instance, particular groups of people are more vulnerable to attacks than others because of their technological skills: an example of a persona singled out not because of suspicion but because of their skills is the system administrator. System administrators form the entry points to communication networks. 'I hunt sys admins' was one of the sentences in the documents, as highlighted in *The Intercept* (Gallagher and Greenwald 2014b). If system administrators are a productive interface for targeting networks, which happened in the Belgacom

73 Interestingly enough, the PEN study of the metaphors used to describe surveillance mentioned above includes a small section (2.2 per cent) of medical metaphors (implanting, infecting, stethoscope) (PEN American Center 2014). It would be interesting to see the development of these tropes over a longer period of time.

case, this raises questions about the exploitation of a particular group of professionals.

6.4.4 Leakage

Table 4 on the next page presents leaky devices. There exist a number of programs that aim to capture 'leaked' metadata or leaked unencrypted internet traffic. Almost every device leaks multiple data in multiple ways.[74] This includes SMS, e-mail, browsing, search queries, android updates, IP addresses, and location data (see table 4).

74 Which does not mean everything is also analysed by the NSA, see Peter Koop's blog (2014).

Leakage		Term of reference or name of the program
Sending (meta)data through phone use	Mobile apps send location to advertising networks Mobile devices send unencrypted text messages, border crossings receive roaming alerts Mobile devices search for cellular messages; registration messages at the border	Happyfoot Dishfire Fast Follower, Happy Foot; JUGGERNAUT
Sending unencrypted internet traffic	Yahoo's unencrypted webcam streams Web searches; unencrypted traffic (calling servers)	OPTIC NERVE (GCHQ) Xkeyscore
Sharing data through updates	Microsoft sends error reports Mobile devices share data through android update	Xkeyscore (+TAO) The mobile surge

Table 4. Examples of leaky devices in the NSA files.
The middle column contains the phenomenon, the right column the associated name of the program, and the left column a summary of the techniques.

A couple of illustrations show how 'leakage' is not merely a by-product but a constitutive part of the life of devices. With respect to mobile phone communication, it is well known that mobile devices reveal their location every time they make a connection to cellular networks. (For details see Soltani and Gellman 2013a.) But a broad range of data is transmitted in a wide variety of instances: 'One secret report showed that just by updating Android software, a user sent more than 500 lines of data about the phone's history and use onto the network' (Larson, Glanz, and Lehren 2014 in *Propublica*). Similarly, with regards to unencrypted internet traffic one often thinks of the registration of web searches or website visits. But data such as webcam imagery,

provided it runs unencrypted over internet cables, can be collected as well. The OPTIC NERVE program of the GCHQ captured data in this way from 1.8 million Yahoo user accounts. Illustrating how some form of 'leakage' is by no means exceptional, the Electronic Frontier Foundation has made an interactive map which gives an impression of which communication can be watched by whom and under which circumstances (Electronic Frontier Foundation 2015).

What we can take from this is that our devices are continuously communicating: they send, search and share. They are seeking and calling other devices. In other words, they lead their own life. Therefore, the question 'how not to leak' might be the wrong question. Or, according to media scholar Wendy Chun (2013), it might be an indication that we are fighting the wrong battle. According to her, computer devices are 'leaky', or 'promiscuous', by default. This use of the term 'promiscuous' originates in computer science and refers to the mode of communication by which data transmission can be freely shared and read. As argued by Chun, computer devices are networked devices and designed to communicate. Locking them into non-promiscuous mode makes them useless. As she describes in her book, and as she demonstrates in her lectures by using simple tools, the suggestion that you are in command of your computer is misleading: 'Using a packet sniffer (...) you can see that your computer constantly wanders without you' (Chun 2006, 3). In one of her lectures (10 October 2013), she argues that it is mainly through branding efforts that these devices have been commonly understood as 'personal devices', which has obfuscated the fact that the tools were promiscuous from the start. As amateur hackers can already read out part of our internet traffic, it is not surprising that a mega-actor such as the NSA gains easy access to these streams of information:

> We have never been in control of our seemingly sealed, yet promiscuous, machines. (...) It is not surprising that things leak, it surprising that we ever imagined our computers not to leak. (...) New media is not about leaks, new media is leak, new media are leaky. (Chun 2013)

6.4.5 Combinations

We should also look at technologies that combine leakage and insertion. The massive use of 'cookies' could be considered a form of bulk insertion conducted by corporate parties such as Google (being massively stored on people's computers), but they are also leaky in that they communicate unencrypted information about the user:

> One secret 2010 British document suggests that the agencies collect such a huge volume of "cookies"—the digital traces left on a mobile device or a computer when a target visits a website—that classified computers were having trouble storing it all. "They are gathered in bulk, and are currently our single largest type of events," the document says. (Larson, Glanz, and Lehren 2014)

Moreover, the NSA also repurposes cookies: cookies leak information about users after which targets can be identified and malware can be placed on their device. This is the 'cookie-hijacking' mentioned in Table 3 and refers to the use of the 'Google PREF cookie' by the NSA.[75] Another example of 'bulk-access' is the use of Microsoft crash reports:

> The automated crash reports are a "neat way" to gain "passive access" to a machine, the presentation continues. Passive access means that, initially, only data the computer sends out into the Internet is captured and saved, but the computer itself is not yet manipulated. Still, even this passive access to error messages provides valuable insights into problems with a targeted person's computer and, thus, information on security holes that might be exploitable for planting malware or spyware on the unwitting victim's computer. (SPIEGEL Staff 2013b)

Also difficult to position, are the widely discussed examples of 'tapping' via the NSA program 'Upstream' and the GCHQ's version 'Tempora'. Splitters are pinned into fiber-optic cables to diverge a volume of the photons that are being transmitted, allowing the agency to receive a copy of the data. This could be considered to be a technology of insertion (in the infrastructure), however, it produces a leakage and captures data thanks to other leaky devices. Therefore, leakage and insertion should not be seen as a strict binary categorisation, but as movements that can overlap and mutually inform, influence and enforce each other.

6.5 Implications for surveillance studies

6.5.1 Questions concerning insertion

Using leakage and insertion as conceptual lenses changes the kind of questions we ask about devices and surveillance practices. The concept of insertion is a necessary complement to existing concepts of surveillance. It expresses more effectively the proximity of surveillance than the detached, more distant notions of the dragnet, harvest or

75 The Electronic Frontier Foundation (2013) explains how this works in 'NSA Turns Cookies (And More) Into Surveillance Beacons'.

assemblage. Jussi Parikka's book *Digital Contagions: a Media Archaeology of Computer Viruses* is inspiring here (Parikka 2007). His work contains a description of how security discourses, both among experts and within the wider political and social imagination, have been influenced by the introduction of viruses, worms and other 'uncontrollable' digital incidents. According to Parikka, the recognition of computer programs as a potential threat can be seen as a posthumanist shift:

> As security discourse moved from emphasizing physical safety to securing the safety of *information patterns*, or internal security, it deterritorialized from practices based on human-to-human interaction to interactions between machines, programs, and the temporal, nonrepresentational processes of digital code. (Parikka 2007, 48)

For Parikka, the virus is also a theoretically interesting object because it allowed new philosophical questions. The virus overflows its boundaries of intentionality (ibid., 5) thus questioning the notion of control.[76] Analogously, what are the questions that the implant raises? With the sudden public presence of implants and injections ('shots') that became visible when the NSA documents were disclosed, will security discourses and practices change as well? As companies and consumers are confronted with new threats, how will these over time translate to public understandings and imaginations? Something seems to happen since major companies have already started with encrypting data.[77] Is the implant of a similar genre as the virus, or do implants and data injections articulate different theoretical repertoires? For instance, as the examples mentioned above show, non-human agency is massively injected into realms associated with intimacy. What does this entail for our notions of surveillance? In other words: What kind of theoretical vocabulary can accommodate the 'data shot'?

There are also implications for Surveillance Studies' preferred 'sites' of research. The infrastructural aspects of surveillance (such as: Who operates the fiber-optic cables?) are obviously important for the future study of the 'materiality of surveillance' (Bauman et al. 2014, 124). Often associated with consumer surveillance (Pridmore 2012; Elmer 2004), cookie-research might need to be placed in a broader

76 The virus becomes uncontrolled and therefore it allows for a theoretical repertoire that resides between Deleuze and Kittler (Parikka 2007, 18). He draws together these two thinkers because with Kittler's work he can sketch a material genealogy of the virus and Deleuze helps understanding the virus as an uncontrollable object in becoming (thus combining concepts that allow us to understand the virus' rigidity and fluidity).

77 For example, Facebook has started collaborating with Open Whisper Systems to provide end-to-end encryption in What'sApp (Open Whisper Systems 2014).

context. Recently Oliver Leistert argued that '[t]he political-economical genealogy of cookies still remains to be written' (2013). That point can now be extended even further: How does the cookie perform its role as an agent of infiltration? Similarly, the range of personas favored by Surveillance Studies might be updated. Considering that system admins are specific targets because of their position in providing access to networks, the figure of the system administrator is in dire need of conceptual attention. And as a side note relating to Parikka's post-humanist shift, this is interesting because it is not just an attack by non-human programs that target other programs and devices, but the use of humans for the infiltration of non-human networks. We should look at what methods system admins use or can use to protect themselves. It is important to tap into already existing modes of interaction with this threat by system admins themselves. Analysis of their (tacit) knowledge and the ways in which system admins see themselves as players in this development would be of great potential value. Lovink and Rossiter (2012) convincingly argue for 'collaborative concept production' with actors within digital culture. Practically, it would be interesting to see to what extent the (social) study of surveillance can move more closely to 'malware research', possibly in collaboration with hackers and security experts.

6.5.2 Questions concerning leakage

With respect to leaky devices, it is productive to follow Chun who states in light of a series of leaks (from leaky social media profiles to the NSA revelations) that it is not only necessary to revisit our perception of control over our 'personal' devices, but that we also need to fundamentally rethink the status of public and private space. Chun's sobering analysis that our devices simply leak implies that we need to rethink what it means habitually to share our lives with these promiscuous devices. If we accept Chun's argument and accept that networked devices are not private property that we can seal off hermetically, but that they are leaky by default, what are the consequences for social life? Chun is not concerned with the NSA primarily, but with an array of practices by which people's intimate data circulate over the web and the harsh and blaming way society responds to this. She pushes us to think about leaky devices not only as threats to privacy but also as devices of publicity and exposure. According to her, they offer a terrain that needs to be claimed: not as safe, private spaces, but as spaces in which people also have the right to take risks and enjoy themselves in public. She compares this conceptualisation of space to

loitering and SlutWalks.[78] According to Chun, if we want to change contemporary practices of shaming and blaming, we should engage with a politics of forgetting and forgiving (versus memory as storage), and possibly, although she primarily considers it to be a societal problem, use technologies of deletion ('technologies that let things die').

Many studies have focused on how mobile technologies enable people to contest, claim and constitute public and social spaces. These studies show how engaging with the production of data through tracking devices allow for specific spatial relations and sociality. This influences the agency that allows people to shape public space (de Souza e Silva and Frith 2010; Sutko and de Souza e Silva 2011; Farman 2014; Tuters and Varnelis 2006). However, the kind of logic that Chun hints at, a logic of deletion, and the spaces she appeals to, inclusive spaces in which one can take risks, are of a particular kind. An appropriate way in which such spaces can be conceptualised in the context of computational data collection is articulated in a recent talk by David Berry, published as a blog post (2014). He elaborates on Hakim Bey's notion of 'opaque temporary autonomous zones', zones that set up evasion or subversion of control. These can be highly visible, yet opaque. Think of squatted cultural centres or occupied squares in the centre of highly surveilled urban areas. Berry stresses the need to think through conditions of possibility of such 'opaque places' for the digital domain. 'These should be fully political spaces, open and inclusive, but nonetheless opaque to the kinds of transparency that computation makes possible' (Berry 2014). This would entail engaging with practices that do not promise or attempt to reach complete privacy or secrecy, but practices that constitute 'shadows', moments of 'hiding in plain sight'. He mentions crypto-practices and the camouflaging of faces to trick surveillance cameras as examples. So the challenge becomes re-inventing repertoires of shadowing and tricking for collective encounters and political purposes:

> we could think about these crypto-practices as (re)creating the possibility of being a crowd, both in the terms of creating a sense of solidarity around the ends of a political/technical endeavour and the means which act as a condition of possibility for it. (Ibid.)

Although many (security) tools and strategies are often considered in terms of individual privacy, reclusive anonymity and private security, it would perhaps be more interesting to follow Berry's line of thought and investigate to what extent these strategies allow for collectives

78 A SlutWalk is a protest march against 'victim-blaming', for instance the often-heard argument that rape or sexual violence against women is caused by their appearance.

to emerge, and in what ways new forms of security and new forms of public space are being constructed. Arguably, one of the clearest examples of how collective presence increases security is Tor, a tool that can be used for anonymous browsing. The more people use Tor, the safer it becomes to be a Tor user. Therefore it is not only a tool that increases the level of anonymity online, but it is also a tool that technically builds on a solidarity principle (Tor 2015). Moreover, the NSA has difficulties with Tor: one of the NSA slides states that the NSA cannot de-anonymise all the Tor users (only fractions) (*The Guardian* 2013a). In fact, the slide offered somewhat of a compliment to Tor's developers by referring to it as '[s]till the King of high secure, low latency Internet Anonymity' (*The Guardian* 2013b).[79] Tor is therefore not only a tool for anonymity for the individual but it also symbolically represents a collective counter-power against NSA surveillance.

Good starting points for a conceptualisation of such strategies that operate on the intersection between crypto, opacity and collectivity can be found in Raley's extensive overview of anti-tracking technologies and both technical and artistic interventions with surveillance (Raley 2013) and Farman's work on the 'creative misuse' of surveillance technologies (Farman 2014). Also, Gehl (2014) offers an analysis of an emerging topic of study that is relevant for these matters. By focusing on social networking on the dark web (e.g. by using Tor), Gehl addresses how in- and exclusion is negotiated in a, technically, anonymous space. According to Gehl, as yet, there is not much scholarship on these 'alternative' forms of sociality (see also Lovink and Rasch 2013). Similar to the Dark Web Social Network there are more examples of technologies that try to accommodate group dynamics in such a way that people become less legible. These practices constitute a promising future topic for investigation. Examples include autonomous servers (e.g. 'Riseup.net'), alternative social networks (N-1), peer-to-peer networks (Retroshare), and alternative internets (GNUnet). Furthermore, there are apps such as InformaCam, which not only aim to increase user-protection: collective use of the app increases its effect.[80] Examples of lightweight techniques that interfere with leaking of browser information 'in public' include browser extensions such as 'Chameleon' (in development, (Ghostwords/chameleon 2015). It detects browser fingerprinting and will give Chrome users the mask of a Tor user (tricking naïve fingerprinters). Tools for obfuscation, such as Ad Nauseam, interfere with leakage by adding randomness by clicking all

79 Although exploiting Firefox vulnerabilities targeted some Tor users, see Schneier's piece in *The Guardian* (2013).

80 InformaCam encrypts but also manages image metadata, thereby facilitating 'event mapping' and digital reporting by using multiple sources. See chapter three in this thesis.

ads (AdNauseam 2005). What these examples share is a recognition of the way in which private and social life is bound to devices and the way in which personal devices are characterised by public leakage. Creative interaction with encryption, deletion and sociability is actively shaping new practices of counter-surveillance.

6.6 Conclusion: Suggestions for future surveillance research

In this chapter, I returned to surveillance studies' revisions of surveillance concepts as was introduced in the first chapter of this thesis (§1.4). On the basis of the NSA disclosures I have distinguished methods for data collection and argued that the NSA files point to interesting directions for surveillance research. I began by describing how the rise of digital networks has made Surveillance Studies move towards a conceptual repertoire of assemblages, profiling and algorithmic sorting. The turn towards assemblages leads to an emphasis on specific sites (such as databases) and specific questions (for instance about sensitive profiles). As a result of this, the tools by which data are captured remain conceptually somewhat unattended. The Snowden archive forces us to also take notice of a wide range of tools with which intelligence agencies collect data. When starting to locate data flows, we encounter things such as cables, servers, implants, injections and leaky devices. In spite of this material dimension, public imagination favours broad immaterial terms to approach surveillance.

With this case study, I tried to draw theoretical consequences from the devices brought to public attention by the disclosures of NSA surveillance and to sketch a different imagination. So, instead of looking at the devices that make surveillance public, I looked at public surveillance devices. I argued that a vocabulary of leakage and insertion addresses both the machinic aspects as well as the intimate, human concerns that arise out of contemporary surveillance practices. The devices, mapped according to two angles, insertion and leakage, point to different lines of movement. The first focuses on devices that get up close and personal and penetrate space that is often seen as private. The massive presence of implants and injections challenges the way in which we use and think of security practices, and there is a need for new theoretical vocabularies to approach these objects. Similarly, for social and humanities scholars of surveillance, this development means that new sites of research in the field of malware are emerging and that collaboration with new partners, system administrators, hackers and security specialists, need to be set up. In addition to sensitive profiles, there is a need to consider sensitive targets such as system admins.

Second, technologies of leaking show the lack of control we have over the devices that we share our lives with and require us to revisit notions of public space. By doing so it becomes worthwhile to focus on ways in which this new reality of surveillance is being subverted. Technologies that combine encryption with sociability might provide the necessary niches for critical public practices to evolve. This constitutes a whole area of interesting input for a future research agenda for the study of material publics in the context of surveillance, which I will elaborate on in the next and last chapter of the thesis.