Chapter 3

3 Putting mass surveillance to use

3.1 Introduction

In this section we look at how the information collected through both passive and active means is put to use by analysts. This chapter describes some of the technical processing of data that determines computers or analysts decisions to analyse data further. This is the moment when the surveillance system can have a major impact on peoples' lives as their digital life is put under scrutiny using further analytics tools.

In this chapter we do not focus on the particular groups being targeted but more generally on the handling of mass surveillance and how this touches on many people who are not specifically targeted.

It is to be expected that most of the people actually targeted by GCHQ and the NSA will be suspected of wrong doing. But deciding on each individual case may not be straightforward, and the secrecy of these processes can easily create indefension. This has been well documented in the case of people put in travel lists, who have little recourse and do not know why this is the case. This was the case of Laura Poitras, the director of the Oscar-winning documentary Citizen Four about Edward Snowden, who had been repeatedly detained at the US border even before the film was made.[i]

Another problem is deciding what constitute legitimate categories of targets. Elsewhere in the report though we describe how GCHQ has variously intruded in the lives of computer engineers and employees of communications companies, people involved in climate negotiations, political activists who use technology (hacktivists), yahoo webcam users, EU commissioners and security experts.

These are only the best documented cases in relation to GCHQ. The NSA has been found monitoring well respected US Muslim leaders in collaboration with the FBI, among others[ii].

This raises another important issue around targeting and how to assess the impacts of mass surveillance. Signal agencies such as the NSA and GCHQ work with other agencies, and much of the original targeting and the outcomes of the surveillance will happen elsewhere. There are no leaked documents from MI5, SIS, or the CIA, where we can see the impacts, such as who is marked for arrest or physical surveillance, and how exactly this is decided.

We know that in the case of the NSA, surveillance can be used for targeted assassination by drone strike, with the target's exact location provided by their mobile phone[iii].

3.1.1 Privacy and power

One of the many arguments used to justify mass surveillance is that the collecting of data in itself does not amount to a breach of privacy. Only if that data is looked at, specifically by a person not a machine, is there a risk. There is also the implication that collected data will only be looked at if there is some suspicion of illegal activity with the oft-repeated line, 'if you have nothing to hide, you have nothing to fear'.

In reality, there are privacy concerns over how our data is stored, analysed and looked at. The processes used will inevitably generate false positives and incorrectly identify individuals as potentially suspect. There are also questions about who has access to this data, which also explored in the following section on sharing data methods.

But the privacy concerns above only cover some of the issues at stake. Novelist John Lanchester was given access to the Snowden files and found that:

"most of what GCHQ does is exactly the kind of thing we all want it to do. It takes an interest in places such as the Horn of Africa, Iran, and North Korea; it takes an interest in energy security, nuclear proliferation, and in state-sponsored computer hacking."

But he also found mass surveillance deeply troubling as it involves a fundamental rebalancing of the power of the state over citizens.[iv]

3.2 The main uses of mass surveillance

3.2.1 Law enforcement and anti-terrorism

Most of the scenarios depicted by defenders of mass surveillance are based on the tactical use of data under the law enforcement model to find terrorists, criminals and child abusers, who are described as 'needles' in the haystack.

But the materials collected will be used for many other purposes, and reducing the debate in this way is not helpful to understand the implications of the technologies described in the previous chapters.

The report by Duncan Campbell for the European Parliament that in the year 2000 publicised the Five Eyes' global system of communications interception ECHELON – a forerunner of the programmes described in this paper – gives some very relevant insights about the use of these kind of surveillance materials that are very relevant to the Snowden leaks.

The report raised concerns about the complicated relationship between law enforcement, national security and Communications Intelligence, which was and is increasingly blurred. Yet maintaining some separation is critical to determine the necessity and proportionality of each programme. Traditional law enforcement involved targeted surveillance or at least a focus on an activity. Intelligence monitoring looks at everything without a predetermined focus to see what patterns or interesting materials appear. The picture that clearly emerges from the leaked documents is that increasingly the latter approach is used for law enforcement.

3.2.2 Strategic Intelligence

This will involve classic scenarios against adversarial states or monitoring global commodities. The recent Snowden leaks confirm this is the case, for example, in relation to Brazilian oil companies.[v] This is not new, the use of ECHELON for economic purposes was well documented at the time of the report, for example in the battle between Boeing and Airbus for contracts in Saudi Arabia. But increasingly, as we can see in leaked documents strategic forecasting involves the direct monitoring of the global populace independently of any state association. This is a new phenomenon.

The ECHELON report made some other important distinctions that would need to be revisited today. The report found that private companies were not able to "task" the systems in order to advance their commercial interests. Governments advance national companies, but this is not the same as allowing them to decide what is targeted. Twenty years ago, the tasking of economic intelligence fell on the Joint Intelligence Committee, the Treasury or the Bank of England. Can we be sure this is still the case with the growing levels of private sector involvement in public affairs.

The second point is that "tasking" information is not the same as disseminating and exploiting it. The decision on the latter was not made by agencies, but by governments. But does this separation still stand, given the vast amounts of information collected and tasked and the increasing use of automated systems?

3.2.3 Reference data

Mass surveillance systems also provide reference data. We saw this in the cases of GCHQ collecting cryptologic data and hacking into companies to map internet infrastructure. The agencies are always planning the next steps building strategic capabilities.

Importantly, all the data that is not flagged up as relating to an individual of interest is used to build the basis for any statistical comparison and other analytics. Understanding anomalies requires an understanding of normal behaviour, and the latter is built through the systematic analysis of the data of regular people.

3.2.4 Cyber attacks and strategic dominance

As we saw in the previous chapters, mass collection of data enables computer network attacks and other cyber operations. In the long run the global system of access to the global network of submarine cables provides the US, UK and allies with strategic dominance of cyber space in the case of armed conflict.

3.2.5 Cyber defense

NSA documents leaked to Der Spiegel show that the agency uses the pervasive monitoring of internet traffic in and out of the country as a kind of digital border to prevent cyberattacks[vi]. It is to be expected that the UK will do the same. This raises the unanswered question of whether the governments of the UK or the US could use this technology implement a kill switch to cut off their countries from much of global communications in an extreme crisis.

3.3 How is bulk data being stored, processed and analysed

3.3.1 Tasking

Automatically, the systems will be looking for people in watch lists of known tasked selectors – emails, phone numbers, social media names – stored in databases such as BROADOAK, but this is only a proportion of the activity. These selectors will have been inputted through the use of tools that generate lists from other sources, in principle all flowing from a policy decision made by government officials outside the agency.

In the case of GCHQ the main tasking tool is called UDAQ,[vii] but little more is known about how it works. There is a highly technical aspect of finding ways of reaching the target or new targets.[viii]

As we saw in the previous section on TEMPORA, these systems apply certain techniques to optimise and minimise the data with tools such as SCISSORS.

From the leaked documents we understand that the rest of the data – relating to millions of people under no suspicion whatsoever – is then processed to separate the content from the "metadata", which is then kept for a further 30 days for further processing. The metadata of innocent people is then analysed for suspicious patterns, or anomalies that can lead to intelligence. From leaked documents relating to XKEYSCORE, this could simply mean the use of anonymous communication software.[ix]

3.3.2 Long term storage

The Interception of Communications Commissioner's Office says that specific information isn't kept for longer than two years unless it meets strict criteria. We do not have GCHQ specific information about how much such data relating to "persons of interest" is kept and for how long.

One important issue is how much content not tasked as part of an investigation but generated through automated searches is kept for analytics, and how this takes place. For example, anyone found to be several degrees of contact with a known terrorist could have their data kept.

In related documents[x] about the US side of the global mass surveillance system XKEYSCORE, it was revealed that the NSA sends some 5% of the data it collects for long term storage in the PINWALE database. This is just one of many NSA databases that run on the basis of content type, or source of where the data came from.

PINWALE can keep the contents for some 5 years, and it's used to mainly store emails but it has apparently been expanded for other forms of content.[xi] PINWALE came to prominence in 2009[xii] after an NSA analyst was apparently investigated for snooping on the personal emails of former president Bill Clinton. At the time GCHQ denied[xiii] they were building a UK version of the system. But these claims are hard to believe given that both countries operate very similar mass surveillance systems.

Other programmes, such as the monitoring of webcams, social media and internal business communications rely on bulk collection technologies to provide the underlying data.

Bulk collection capabilities allow for the development of many projects. German news organisation Der Spiegel revealed that the NSA conducts extensive mass surveillance of financial transactions, collecting up to 180 million datasets by 2011 with the help of GCHQ. The so-called “Follow the Money” programme apparently led GCHQ’s lawyers to raise concerns about the collection, storage and sharing of such “politically sensitive" and "bulk data – rich personal information. A lot of it is not about our targets."

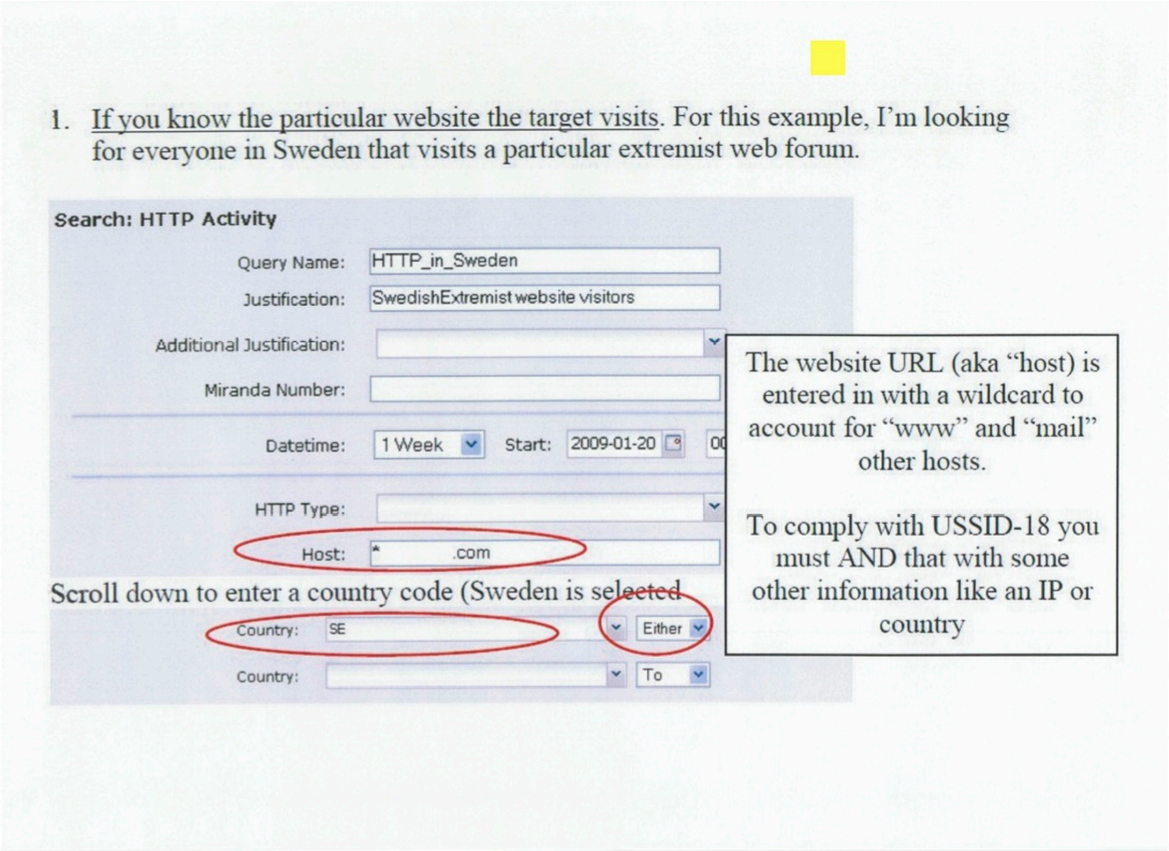
3.3.3 Analysis

The analysts have to build on those initial selectors – called "seeds" - to identify who they belong to, find more information about known targets and find new ones. This process tends to involve finding more selectors that can then be fed back into the system for further searches. The ultimate aim is to build a complete picture of activities and relationships that can be useful. This is another process that appears to have been hugely transformed by the use of computers and automation.

The data of all internet users is subjected to experimentation to try to find patterns and new insights with the use of XKEYSCORE and other ancillary tools. The NSA put it in terms of "shifting their analytic approach from a production to a discovery bias".[xiv]

Besides the initial mass collection of the data, this is one of the most worrying aspects of the system. It is unclear what happens if your information is flagged as suspicious, but it is very possible that a permanent marker is placed and related data is kept longer than 30 days.

Searches with XKEYSCORE can be very broad, for example, "everyone in Sweden who visits a particular web forum”.[xv] The justification in that case is simply "SwedishExtemist website visitor", which suggests that justifications can be very flexible and adapted to the need of the analyst.



1. If you know the particular website the target visits. For this example, I'm looking for everyone in Sweden that visits a particular extremist web forum.

Search: HTTP Activity

Query Name:	HTTP_in_Sweden
Justification:	SwedishExtremist website visitors
Additional Justification:	
Miranda Number:	
Datetime:	1 Week Start: 2009-01-20 00
HTTP Type:	
Host:	* .com

Scroll down to enter a country code (Sweden is selected

Country:	SE	Either
Country:		To

The website URL (aka "host) is entered in with a wildcard to account for "www" and "mail" other hosts.

To comply with USSID-18 you must AND that with some other information like an IP or country

The analyst will also have to enter a Miranda number relating to a particular need for information, and possibly some additional justification from a drop down menu. No additional warrants are required if a GCHQ analyst believes that an individual is abroad and s/he will be covered by the broad certificates allowing international surveillance. At least in the US, searches by analysts are only periodically reviewed by supervisors[xvi]. Operatives can also use the tool for "tasking" new targets; ordering future focused monitoring of individuals or communications categories.

The US National Academy of Science has published a report on bulk surveillance[xvii] that gives some examples of the activities involved in SIGDEV. Contact chaining would involve discovering possible associations through the matching of many contact sources (address books, social media etc) or location proximity to a target. An analyst could also try to discover alternate identifiers for a known target, such as different email addresses, etc. Triage involves categorising identifiers according to the danger they may represent. This is a very important activity with important rights implications as false positives can have have very serious consequences. It involves looking for connections to events and people.

The targeting of innocent people in bulk surveillance systems appears to be widespread as programmes try to build connections or individuals become useful in order to reach a target.

The actual XKEYSCORE software has been leaked, showing that it tracks broad categories of people, including thousands of users of anonymisation software[xviii].

The NSA – under the operation codenamed AURORAGOLD[xix] – spies on hundreds of companies and organisations, including the UK based trade group GSM Association, in an effort to find security weaknesses in cell phone technology that it can exploit for surveillance. This has led them to include over emails of innocent employees of these organisations in their tasking systems. It is likely that many of these are targeted by the UK as well, if they are based overseas.

Indeed, one of the key issues with bulk surveillance is that it generates lots of potential leads. And as we have seen in recent cases, such as the Charlie Hebdo and Woolwich killings, the capacity to act downstream by security forces is a lot more limited. The agencies have many strategies to increasing the technical capacity to handle large datasets and reducing the amount of data.

The NSA has built a big data platform called GHOSTMACHINE[xx] that provides analytics capabilities for handling the huge numbers generated in bulk collection.

Triage and many other activities require reducing the number of targets. GCHQ and the NSA used a system called ECHOBASE[xxi] that uses automation to bring down the numbers generated by other tools to a level where they can be analysed manually.

The report by the US National Academy of Science cautiously proposes some possible technical controls to restrict access to data in order to give citizens assurances against abuses, including automatically restricting the types of queries to access bulk data. They also propose automating the auditing of the queries used by analysts, which would allow for more external scrutiny without compromising the sensitive details of the search queries.

The authors stress the caveat that nobody outside the agencies fully understands their internal processes. Indeed this is a fundamental obstacle to building any trust. Any technical measures may not be sufficient without changes to the wider practices and attitudes towards transparency.

3.3.4 Dissemination

This is the aspect that has received less attention in the debates raised by the Snowden revelations. Partly, this has to do with a lack of information about the UK in comparison with the US in the leaked documents. In addition the sensitivity of the final outputs of intelligence would mean that either Snowden didn’t have access or those with access to the documents apply extra caution.

The Guardian newspaper gave some information from a direct source at the time of the first revelations of TEMPORA[xxii]:

"The data collected provides a powerful tool in the hands of the security agencies, enabling them to sift for evidence of serious crime. According to the source, it has allowed them to discover new techniques used by terrorists to avoid security checks and to identify terrorists planning atrocities. It has also been used against child exploitation networks and in the field of cyberdefence.

It was claimed on Friday that it directly led to the arrest and imprisonment of a cell in the Midlands who were planning co-ordinated attacks; to the arrest of five Luton-based individuals preparing acts of terror, and to the arrest of three London-based people planning attacks prior to the Olympics."

There is little information in the Snowden files on any direct outcomes of mass surveillance in the UK in terms of intelligence reports. Leaked documents have shown the contributions that mass surveillance programmes have made to presidential briefings in the US. We have heard that some 300 terrorists have been captured, and there has been some information leaked about the use of PRISM during the Olympics.

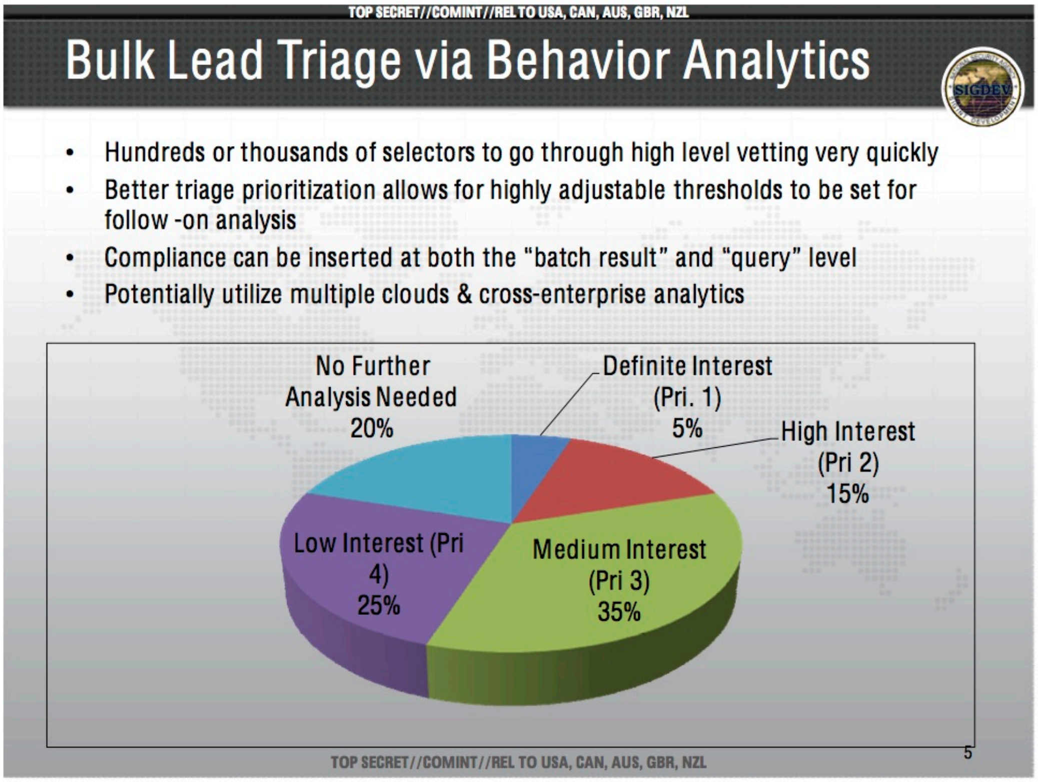
The kind of information gleaned by the NSA or GCHQ would in most cases be fed to other intelligence or law enforcement bodies, who would then try to incorporate it into their work. However, it seems likely that increasingly tasks that were carried out by human operatives in other agencies are now performed by computers in signals organisations. For example, the tradecraft of "gisting" - writing snappy summaries – could be partially replaced by machines.

3.4 Additional issues with processing

3.4.1 Predictive analytics and profiling

We often hear that bulk collection is necessary in order to find the "needle in the haystack". Sir Iain Lobban, head of GCHQ from 2008 until January 2014, used these words extensively in his appearance in Parliament[xxiii] in November 2013. But this is a misleading argument, because modern profiling systems do not operate with such clear categories. Instead everyone is analysed and subjected to dynamic risk analyses that treat every innocent citizen as a potential suspect. There is rarely absolute certainty on who may be a needle.

The ECHOBASE tool we mentioned above categorises potential leads already flagged elsewhere, but it does not give simple answers:



TOP SECRET//COMINT//REL TO USA, CAN, AUS, GBR, NZL

Bulk Lead Triage via Behavior Analytics

- Hundreds or thousands of selectors to go through high level vetting very quickly
- Better triage prioritization allows for highly adjustable thresholds to be set for follow -on analysis
- Compliance can be inserted at both the "batch result" and "query" level
- Potentially utilize multiple clouds & cross-enterprise analytics

No Further Analysis Needed 20%
Definite Interest (Pri. 1) 5%
High Interest (Pri 2) 15%
Medium Interest (Pri 3) 35%
Low Interest (Pri 4) 25%
SIGDEV
5

TOP SECRET//COMINT//REL TO USA, CAN, AUS, GBR, NZL

The details of most systems used by GCHQ remain mostly secret, but scholars such as Louise Amoore have been studying profiling systems in other areas of national security such as borders and detention orders.[xxiv] These systems are based on data mining techniques that have been developed in commercial applications at casinos, in fraud detection, etc.

Border controls are a good example of security profiling. Here we could imagine actual lists of dangerous people who will be prevented from entering the country or possibly taking a plane. The Home Office describes the National Border Targeting Centre (NBTC) – responsible for border profiling – in these simple terms:

"More than 100 million passenger movements in and out of the UK were checked against UK Border Agency and police watch lists last year. (Home Office press release)"[xxv]

But this does not fully reflect the way the newer border controls operate for everyone. There are indeed secret watch lists with named individuals, which raise their own questions about accountability and due process. However, the computers at NBTC perform more complex real-time risk assessments of all airline passengers, where those with a substantial risk score are flagged when they cross the border.

This score is not fixed but will be constantly recalculated. Buying a ticket in cash, or having taken a previous trip to a troublesome country, when combined with other factors could trigger an alert.

Each individual element of a security risk alert may be completely lawful innocent behaviour. This has been evidenced in security deportation orders. In the words of a defence lawyer at one such case at the Special Immigration Appeals Commission:

“Neither we nor our clients were given the 'ingredients' of the mosaic – we were only given conclusions, expressed in the form 'we assess that X has been involved in attack planning.' This is the way it operates, piecing together fragments which in themselves are innocent.”[xxvi]

Another complication for the narrative of the needle in the haystack are the new risk models developed since the War on Terror. There is a wealth of research on how the unimaginable events of 911 fundamentally changed security risk calculations.[xxvii]

The new basic premise is that the mere possibility of a high impact event should be enough to trigger a response. The UK National Security Strategy encapsulates this fatalistic way of thinking:

“(...) this strategy must allow the Government to make choices about the risks we face. Of course, in an age of uncertainty the unexpected will happen...“[xxviii]

The fixed categories of innocent and suspect become more blurred when we are not trying to establish probabilities but simply possibilities. For example, in the period around Christmas 2012-13 NBTC issued 4,900 alerts to border agencies, and these carried out 237 arrests.[xxix]

The sophisticated mass surveillance programs of the NSA and GCHQ treat all Internet users like international airplane passengers. After the needles are found, the haystack is not discarded but used to build databases of evolving risks. For example, the lawful use of encryption in emails or searches of the TOR website will increase your risk score.[xxx]

In the context of air travel or border controls there could be an argument that this intrusion is necessary for the safety of the flight, and limited to that context. But mass surveillance programs are permanently tracking what we do online.

In fact, the New York Times newspaper reported that the NSA actually monitors airline information data as part of their profiling and analysis of social networks, including of American citizens [xxxi].

3.4.2 Machine processing

Advances in machine learning and artificial intelligence should make us question the current focus on human activities in surveillance legislation and policy. The often-heard argument that there is no mass surveillance if “nobody reads, looks or listens” to the collected information is out of touch with the capacities of modern digital systems.

Sir Iain Lobban, head of GCHQ from 2008 until January 2014, made a big point about this, when he stressed in evidence to Parliament[xxxii] that operatives do not access all the collected data:

"We do not spend our time listening to the telephone calls or reading the e-mails of the majority, the vast majority that would not be proportionate. It would not be legal. We do not do it."

But computer analysis is just another kind of intrusion on privacy. And the agencies have extensive capabilities with little external regulation.

The need to read emails is reduced by technology that can process text content. This is not even top secret. The NSA makes available under several tools to process natural language texts its technology transfer program. These do not require a human operative to actually read them.[xxxiii] The NSA also licenses technologies for handling voice, including speaker recognition. It is to be expected that GCHQ uses these or similar technologies:

“NSA's acoustic technologies include methods for identification, extraction, and analysis of voice and voice signals. Additional technologies include foreign language voice recognition, duplicate voice identification, and methods of measuring voice enhancement.” [xxxiv]

In recent years facial recognition has finally reached maturity. Despite the lack of publicly available information on their operational effectiveness in the field[xxxv] these systems are being rolled into production everywhere.

There are documents showing that GCHQ engages in computerised facial recognition. As we showed above, GCHQ has tapped into the private webcam communications of innocent Yahoo! Subscribers. In its own documents, the agency explains it did this as an experiment to improve facial recognition.

Facial recognition and text analysis are not even the most advanced technologies. State of the art machine learning tools are capable not just of recognising a specific individual's face, but of learning to classify faces based on attributes such as hair style or expression.[xxxvi] Computer systems from Google can teach themselves new concepts from scratch, by looking at pictures and videos, such as figuring out what is a cat.[xxxvii]

These development have far-reaching implications for regulating surveillance. Claims that intrusion only takes place when humans are involved in “reading, listening to, or looking at” are hard to sustain, give the information that can be gleaned by computers alone.

3.4.3 Social Media monitoring

NBC news has reported that GCHQ uses a variety of tools for their real time analysis of social media, including commercially available software.[xxxviii] But it is the capacity to intercept and collect bulk raw Internet data that makes it possible in the first place. The SQUEAKY DOLPHIN programme is nominally focused on the general analysis of social media trends; the reports mention protests in Bahrain as one the scenarios used in the project.

But according to NBC, there are documents that show that GCHQ also uses Twitter data to identify specific users around the world.

The analysis appears to have looked mainly at social media sources that were not encrypted, such as Twitter, Facebook, Blogger and Youtube. Since the reports were published several of these services have started encrypting their traffic.

In addition it has been reported that GCHQ carries out targeted automated monitoring of known social media accounts. The case relates to known security experts being targeted under the project LOVELY HORSE.[xxxix] In such cases it seems likely that a targeted warrant should be used.

3.5 Conclusion

Profiling techniques are powerful and can easily be misdirected. It is hard to justify profiling everyone, but that is in effect what is happening. The systems provide constant risk analysis of each citizen whose data is contained in the database.

Machine processing is a major threat to citizens' privacy. Arguments that humans do not examine material are inadequate, when compared with the power of computers to analyse and sort.

Of course private data is not the only source of information for GCHQ, who harvest public information as well. Ethical issues occur with both public and private sources of data. Just because personal information is public, it does not mean that anyone has permission to do anything they like with it. This is especially true for nation states, because of the power they wield. The very least we deserve is an honest debate about how and when agencies can use public data.

The chapter again saw the same high levels of integration of US and UK operations. The implications of this are outlined in Chapter 8, which looks at all of the threats emerging from this picture.

[i] http://www.salon.com/2012/04/08/u_s_filmmaker_repeatedly_detained_at_border/

[ii] https://firstlook.org/theintercept/2014/07/09/under-surveillance/

[iii] https://firstlook.org/theintercept/2014/02/10/the-nsas-secret-role/

[iv] http://www.theguardian.com/world/2013/oct/03/edward-snowden-files-john-lanchester

[v] http://www.theguardian.com/world/2013/sep/09/nsa-spying-brazil-oil-petrobras

[vi] http://www.spiegel.de/international/world/new-snowden-docs-indicate-scope-of-nsa-preparations-for-cyber-battle-a-1013409-2.html

[vii] https://snowdenarchive.cjfe.org/greenstone/collect/snowden1/index/assoc/HASH019b/107887fc.dir/doc.pdf

[viii] http://www.spiegel.de/media/media-35551.pdf

[ix] http://daserste.ndr.de/panorama/aktuell/nsa230_page-1.html

[x] https://www.documentcloud.org/documents/894406-nsa-slides-xkeyscore.html

[xi] https://en.wikipedia.org/wiki/

[xii] http://www.wired.com/2009/06/pinwale

[xiii] http://www.theguardian.com/technology/blog/2009/jun/18/nsa-pinwale-email-snooping

[xiv] http://cryptome.org/2013/11/nsa-sigint-strategy-2012-2016.pdf

[xv] http://cryptome.org/2013/12/nsa-se-fra-xkeyscore-slide.pdf

[xvi] http://www.theguardian.com/world/2013/jul/31/nsa-top-secret-program-online-data

[xvii] http://www.dni.gov/index.php/newsroom/press-releases/210-press-releases-2015/1161-national-academy-of-sciences-releases-ppd-28-report-bulk-collection-of-signals-intelligence-technical-options

[xviii] http://daserste.ndr.de/panorama/aktuell/nsa230_page-1.html

[xix] https://firstlook.org/theintercept/2014/12/04/nsa-auroragold-hack-cellphones

[xx] https://www.aclu.org/files/natsec/nsa/20140130/2013.12.10%20Ghost%20Machine.pdf

[xxi] https://www.aclu.org/files/natsec/nsa/ghostmachine-identifier-lead-triage-with-echobase.pdf

[xxii] http://www.theguardian.com/uk/2013/jun/21/gchq-cables-secret-world-communications-nsa

[xxiii] http://www.bbc.co.uk/news/uk-politics-24848186

[xxiv] Amoore, L. (2013). *The Politics of Possibility. Risk and Security Beyond Probability*. Duke University Press.

[xxv] http://www.mynewsdesk.com/uk/pressreleases/home-office-government-ramps-up-passenger-screening-382694

[xxvi] Interview with Louise Amoore, in above.

[xxvii] Amoore, L., & de Goede, M. (2008). *Risk and the War on Terror*. Taylor & Francis.

[xxviii] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/61936/national-security-strategy.pdf

[xxix] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/311126/PartnerBulletinFebruary.pdf

[xxx] http://www.theguardian.com/world/interactive/2013/jun/20/exhibit-b-nsa-procedures-document

[xxxi] http://www.nytimes.com/2013/09/29/us/nsa-examines-social-networks-of-us-citizens.html?pagewanted=all&_r=1

[xxxii] http://www.bbc.co.uk/news/uk-politics-24848186

[xxxiii] https://www.nsa.gov/research/_files/tech_transfers/nsa_technology_transfer_program.pdf

[xxxiv] https://www.nsa.gov/research/tnw/tnw193/article2.shtml

[xxxv] https://www.nyu.edu/projects/nissenbaum/papers/facial_recognition_report.pdf

[xxxvi] https://www.facebook.com/publications/225061261024135/

[xxxvii] http://googleblog.blogspot.co.uk/2012/06/using-large-scale-brain-simulations-for.html

[xxxviii] http://investigations.nbcnews.com/_news/2014/01/27/22469304-snowden-docs-reveal-british-spies-snooped-on-youtube-and-facebook?lite

[xxxix] http://cryptome.org/2015/02/gchq-lovely-horse-intercept-15-0204.pdf