http://www.guardian.co.uk/uk/2013/jun/21/gchq-mastering-the-internet/print

Mastering the internet: how GCHQ set out to spy on the World Wide Web

Project Tempora – the evolution of a secret programme to capture vast amounts of web and phone data

Ewen MacAskill, Julian Borger, Nick Hopkins, Nick Davies and James Ball

6/21/2013

The memo was finished at 9.32am on Tuesday 19 May 2009, and was written jointly by the director in charge of GCHQ's top-secret Mastering the Internet (MTI) project and a senior member of the agency's cyber-defence team.

The internal email, seen by the Guardian, was a "prioritisation and tasking initiative" to another senior member of staff about the problems facing GCHQ during a period when technology seemed to be racing ahead of the intelligence community.

The authors wanted new ideas – and fast.

"It is becoming increasingly difficult for GCHQ to acquire the rich sources of traffic needed to enable our support to partners within HMG [Her Majesty's government], the armed forces, and overseas," they wrote.

"The rapid development of different technologies, types of traffic, service providers and networks, and the growth in sheer volumes that accompany particularly the expansion and use of

the internet, present an unprecedented challenge to the success of GCHQ's mission. Critically we are not currently able to prioritise and task the increasing range and scale of our accesses at the pace, or with the coherence demanded of the internet age: potentially available data is not accessed, potential benefit for HMG is not delivered."

The memo continued: "We would like you to lead a small team to fully define this shortfall in tasking capability [and] identify all the necessary changes needed to rectify it." The two chiefs said they wanted an initial report within a month, and every month thereafter, on what one of them described as "potential quick-win solutions not currently within existing programme plans".

Though this document only offers a snapshot, at the time it was written four years ago some senior officials at GCHQ were clearly anxious about the future, and casting around for ideas among senior colleagues about how to address a variety of problems.

According to the papers leaked by the US National Security Agency (NSA) whistleblower Edward Snowden, Cheltenham's overarching project to "master the internet" was under way, but one of its core programmes, Tempora, was still being tested and developed, and the agency's principal customers, the government, MI5 and MI6, remained hungry for more and better-quality information.

There was America's NSA to consider too. The Americans had been pushing GCHQ to provide better intelligence to support Nato's military effort in Afghanistan; a reflection, perhaps, of wider US frustration that information sharing between the US and the UK had become too lopsided over the past 20 years.

In the joint instruction from 2009, the director had twice mentioned the necessity to fulfil GCHQ's "mission", but the academics and commentators who follow Britain's intelligence agencies are unsure exactly what this means, and politicians rarely try to define it in any detail.

The "mission" has certainly changed and the agency, currently run by Sir Iain Lobban, may be under more pressure now than it has ever been.

The issues, and the enemies, have become more complex, and are quite different from the comparatively simple world into which Britain's secret services were born in 1909.

At the time, concern about German spies living in the UK led to the establishment of a Secret Service Bureau and, at the start of the first world war, two embryonic security organisations began to focus on "signals intelligence" (Sigint), which remains at heart of GCHQ's work.

The codebreakers of Bletchley Park became heroes of the second world war as they mastered the encryption systems used by the Nazis. And the priority during the cold war was Moscow.

During these periods GCHQ's focus was clear, and the priorities of the "mission" easier to establish.

There was no parliamentary scrutiny of its work so the agency, which moved from Milton Keynes to Cheltenham in the early 1950s, existed in a peculiar limbo.

That changed, and with it the boundaries of its work, with the 1994 Intelligence Services Act (Isa), which gave a legal underpinning to the agency for the first time. The act kept the powers and objectives of GCHQ broad and vague.

The agency was tasked with working "in the interests of national security, with particular reference to the defence and foreign policies of Her Majesty's government; in the interests of the economic wellbeing of the United Kingdom; and in support of the prevention and the detection of serious crime".

Reviewing the legislation at the time, the human rights lawyer John Wadham, then legal director of Liberty, highlighted the ambiguities of the expressions used, and warned that the lack of clarity would cause problems and concern.

"National security is used without further definition. It is true the courts themselves have found it impossible to decide what is or what is not in the interests of national security. The reality is that 'national security' can mean whatever the government of the day chooses it to mean." The same could be said for the clause referring to "economic wellbeing".

Arguably, GCHQ's responsibilities have broadened even further over the past decade: it has become the UK's lead agency for cyber-security – identifying the hackers, criminal gangs and state actors who are stealing ideas, information and blueprints from British firms.

Alarmed by the increase in these cyber-attacks, and faced with billions of pounds' worth of intellectual property being stolen every year, the government made the issue a tier-one priority in the 2010 strategic defence and security review. In a time of cuts across Whitehall, the coalition found an extra £650m for cyber-security initiatives, and more than half was given to GCHQ. It has left the agency with a vast array of responsibilities, which were set out in a pithy internal GCHQ memo dated October 2011: "[Our] targets boil down to diplomatic/military/commercial targets/terrorists/organised criminals and e-crime/cyber actors".

All this has taken place during an era in which it has become harder, the intelligence community claims, for analysts to access the information they believe they need. The exponential growth in the number of mobile phone users during the noughties, and the rise of a new breed of independent-minded internet service providers, conspired to make their work more difficult, particularly as many of the new firms were based abroad, outside the jurisdiction of British law.

Struggling to cope with increased demands, a more complex environment, and working within laws that critics say are hopelessly outdated, GCHQ starting casting around for new, innovative ideas. Tempora was one of them.

Though the documents are not explicit, it seems the Mastering the Internet programme began life in early 2007 and, a year later, work began on an experimental research project, run out of GCHQ's outpost at Bude in Cornwall.

Its aim was to establish the practical uses of an "internet buffer", the first of which was referred to as CPC, or Cheltenham Processing Centre.

By March 2010, analysts from the NSA had been allowed some preliminary access to the project, which, at the time, appears to have been codenamed TINT, and was being referred to in official documents as a "joint GCHQ/NSA research initiative".

TINT, the documents explain, "uniquely allows retrospective analysis for attribution" – a storage system of sorts, which allowed analysts to capture traffic on the internet and then review it.

The papers seen by the Guardian make clear that at some point – it is not clear when – GCHQ began to plug into the cables that carry internet traffic into and out of the country, and garner material in a process repeatedly referred to as SSE. This is thought to mean special source exploitation.

The capability, which was authorised by legal warrants, gave GCHQ access to a vast amount of raw information, and the TINT programme a potential way of being able to store it.

A year after the plaintive email asking for new ideas, GCHQ reported significant progress on a number of fronts.

One document described how there were 2 billion users of the internet worldwide, how Facebook had more than 400 million regular users and how there had been a 600% growth in mobile internet traffic the year before. "But we are starting to 'master' the internet," the author claimed. "And our current capability is quite impressive."

The report said the UK now had the "biggest internet access in Five Eyes" – the group of intelligence organisations from the US, UK, Canada, New Zealand and Australia. "We are in the golden age," the report added.

There were caveats. The paper warned that American internet service providers were moving to Malaysia and India, and the NSA was "buying up real estate in these places".

"We won't see this traffic crossing the UK. Oh dear," the author said. He suggested Britain should do the same and play the "US at [their] own game … and buy facilities overseas".

GCHQ's mid-year 2010-11 review revealed another startling fact about Mastering the Internet.

"MTI delivered the next big step in the access, processing and storage journey, hitting a new high of more than 39bn events in a 24-hour period, dramatically increasing our capability to produce unique intelligence from our targets' use of the internet and made major contributions to recent operations."

This appears to suggest GCHQ had managed to record 39bn separate pieces of information during a single day. The report noted there had been "encouraging innovation across all of GCHQ".

The NSA remarked on the success of GCHQ in a "Joint Collaboration Activity" report in February 2011. In a startling admission, it said Cheltenham now "produces larger amounts of metadata collection than the NSA", metadata being the bare details of calls made and messages sent rather than the content within them.

The close working relationship between the two agencies was underlined later in the document, with a suggestion that this was a necessity to process such a vast amount of raw information.

"GCHQ analysts effectively exploit NSA metadata for intelligence production, target development/discovery purposes," the report explained.

"NSA analysts effectively exploit GCHQ metadata for intelligence production, target development/discovery purposes. GCHQ and NSA avoid processing the same data twice and proactively seek to converge technical solutions and processing architectures."

The documents appear to suggest the two agencies had come to rely on each other; with Tempora's "buffering capability", and Britain's access to the cables that carry internet traffic in and out of the country, GCHQ has been able to collect and store a huge amount of information.

The NSA, however, had provided GCHQ with the tools necessary to sift through the data and get value from it.

By May last year, the volume of information available to them grew again, with GCHQ reporting that it now had "internet buffering" capability running from its headquarters in Cheltenham, its station in Bude, and a location abroad, which the Guardian will not identify. The programme was now capable of collecting, a memo explained with excited understatement, "a lot of data!"

Referring to Tempora's "deep dive capability", it explained: "It builds upon the success of the TINT experiment and will provide a vital unique capability.

"This gives over 300 GCHQ and 250 NSA analysts access to huge amounts of data to support the target discovery mission. The MTI programme would like to say a big thanks to everyone who has made this possible … a true collaborative effort!"

Tempora, the document said, had shown that "every area of ops can get real benefit from this capability, especially for target discovery and target development".

But while the ingenuity of the Tempora programme is not in doubt, its existence may trouble anyone who sends and receives an email, or makes an internet phone call, or posts a message on a social media site, and expects the communication to remain private.

Campaigners and human rights lawyers will doubtless want to know how Britain's laws have been applied to allow this vast collection of data. They will ask questions about the oversight of the programme by ministers, MPs and the intelligence interception commissioner, none of whom have spoken in public about it.