Watching him watching you:
The latest surveillance technologies have raised fears of a Big Brother style society

By Eugenie Samuel, New Scientist, June 14, 2013

PICTURE this. You work in the second tallest skyscraper in Seattle. Your partner recently bought a load of chemical fertilizer for your rose garden. A friend of yours, with whom you're always talking on the phone, is writing a paper on Middle Eastern politics and has spent time on a website that is sometimes visited by Palestinian extremists. It's all completely innocent. But the Pentagon might not see it that way. They find the coincidence of your friend's online activity and the shipment of explosive material to your address suspicious and haul you in for questioning. The next thing you know, your boss is told that you are a potential terror suspect and should not be let into the building.

A paranoid fantasy? Not according to the US Congress. The Pentagon's research agency, DARPA, has been working on a multimillion-dollar project that aims to develop tools to "mine" intelligence files and publicly available data for suspicious patterns in credit card bills, phone records, airline reservations, online news reports and other sources, allowing intelligence gathering on an unprecedented scale. The plan is to develop several systems that could identify terrorists and alert the authorities before they strike. But critics of the project are calling for it to be scrapped, saying it allows the US government an unprecedented window on every citizen's life, and risks causing miscarriages of justice.

When the Total Information Awareness (TIA) project was announced in January, the media quickly--and inevitably--shouted its concerns about an Orwellian Big Brother. But, argued the Pentagon, if you have nothing to hide, then you have everything to gain. Take the case of Mohammed Atta, the alleged ringleader of the 11 September2001 attacks. Twelve days before the attacks, Atta logged onto the web and used his credit card to buy two one-way air tickets. The tickets were for [travel](http://go.galegroup.com/ps/retrieve.do?sgHitCountType=None&sort=RELEVANCE&inPS=true&prodId=GPS&userGroupName=mayf29809&tabID=T002&searchId=R2&resultListType=RESULT_LIST&contentSegment=&searchType=BasicSearchForm&currentPosition=17&contentSet=GALE%7CA104080930&&docId=GALE|A104080930&docType=GALE&role=AONE) on the same day as 17 fellow Al-Qaida members. If the transaction had been spotted it would have set alarm bells ringing, but unfortunately the technology to do so was not in place. However, that is exactly the kind of clue that TIA could pick up in the future.

Despite DARPA's good intentions, it still needs to prove TIA will not invade people's privacy. To persuade Congress to allow research on TIA to continue, DARPA produced a report last month that explains exactly how they plan to protect people's privacy. Unlike DARPA's first documents on TIA, which carried a large unblinking eye as the logo, the new report sports an abstract pattern of three overlapping circles and a swirl. Not only that, but at the end of a week when the US had been on high terrorist alert, TIA suddenly stood for [Terrorism](http://go.galegroup.com/ps/retrieve.do?sgHitCountType=None&sort=RELEVANCE&inPS=true&prodId=GPS&userGroupName=mayf29809&tabID=T002&searchId=R2&resultListType=RESULT_LIST&contentSegment=&searchType=BasicSearchForm&currentPosition=17&contentSet=GALE%7CA104080930&&docId=GALE|A104080930&docType=GALE&role=AONE) Information Awareness.

It's tempting to dismiss this as a cynical act of opportunistic spin, capitalising on fear while playing down the intrusive nature of the project. That's certainly how privacy advocates are seeing it (New Scientist, 31 May, p 8). But maybe they aren't looking at the bigger picture. Spotting terrorists before they strike will save lives, so there is a need for surveillance technologies like TIA. However, that needn't mean intruding on the lives of innocent people. What's missing is something to keep an eye on TIA and make sure no one misuses or abuses it--an electronic guardian, in other words. In last month's report, DARPA builds on a study commissioned last year into technologies that could protect privacy in the context of data mining. The study concluded that there is potential for electronic guardians--if only we find the will to develop them.

It seems that the public's reaction to DARPA's original proposal is encouraging the agency to find that will. Many people felt that DARPA's actions might lead to persecution, as it did in the 1950s when senator Joseph McCarthy used people's tax files and other records to persecute government dissenters in the name of anticommunism. But, as DARPA has now been forced to show, it is willing to look into technologies that can keep an eye on other, more worrisome technologies. Thanks to Congress, and public outcry, it may be that TIA can produce exactly what the 21st century needs: effective surveillance that operates under cast-iron accountability.

Although TIA includes several different technologies--human recognition at a distance; activity recognition, such as planting bombs; and automatic translation of languages--the main thrust is software "bots". These are designed to search governmental or commercial databases, including lists of credit card transactions and phone calls. The bots work on a similar principle to search engines, such as Google, or the US's Echelon system, which intercepts international communications and searches for terrorist-related terms.

Unlike Google's blunt searches, the aim of TIA is to develop bots that look for characteristic and very specific patterns of terrorist activity. "TIA is not data mining, it is hypothesis driven," says Jan Walker, spokeswoman for DARPA. To do this, TIA bots will upload themselves to other locations and incorporate themselves into the host code. They might reorganise a database of credit card records that is only searchable under credit card number and name, so it can be searched for certain kinds of purchase as well. Once on the host system, the bot can run its search there, pulling out complex patterns in the data, such as the coincidence of certain phone calls or travel plans. This information is then returned to the intelligence agency via the internet or other data links.

But there is a way to stop government agents running indiscriminate, overly blunt searches. You need the database that is being searched to analyze the bot, before it runs its program, and decide whether or not to allow it to continue. However, getting a database to protect itself isn't as straightforward as you might think; indeed it's been recognized as a huge challenge ever since British mathematician Alan Turing came up with his "halting problem" in 1936. This says that it is impossible to write a program able to tell whether another program terminates, or whether it runs forever, without running the other program. Just as it is impossible to find out whether a program terminates, it is also impossible to find out what it does without running it. This means there is no way for a database code to tell what an intrusive TIA bat might do with its data if it were allowed to run.

George Necula of the University of California at Berkeley and Peter Lee of Carnegie Mellon University in Pittsburgh, Pennsylvania, reckon they have found a way round the problem. "We put the burden on the incoming code to satisfy the rules, not on the host system to assess it," Necula says. Under his scheme, developed in 1997, a bot must contain a subroutine called a "proof" that explains how it works. A well-protected database can inspect the program to see whether it has the structure the proof describes. It's rather like trying to find your way through a maze, which is impossible to do without going in and exploring. But if someone gives you a path through the maze, it's a relatively simple matter to inspect the maze and check that what they say is true.

Necula showed that if a host system demands sufficient proof, for example, of enough of the possible paths the bot wants to take, it could find out enough about the workings of the bot to answer most security and privacy concerns. He has even designed schemes that make it possible for host systems to tell whether incoming bots were well-written and well-behaved programs, or whether they were sketchy or viral code that could cause damage.

There is no reason, says Necula, why his schemes could not be adapted to the problem of protecting privacy. Suppose a database contains private information on the itemized phone bills of thousands of people. The host could be designed to let in bots that intend to search for patterns of behavior--repeated calls to certain numbers for example. But it could refuse bots that are liable to be too nosy, pulling out names and addresses of people who do not fit into the patterns. The bot could even be restricted to pulling out patterns without people--flagging the presence of suspicious travel plans, for example, without the hot knowing whose they are. With such schemes in place, intelligence agencies could find out if there is relevant information in a particular database, without rifling through the private lives of everyone in it. They could then apply to the courts for a defined search warrant to let them have better access to particular aspects of the data. It's the kind of process that, according to last month's report, DARPA is committed to developing.

Spotting the snoopers

The report also claims that TIA will develop technology to monitor its own use. "Among the controls being researched are automatic audit trails to document who accesses the system and how it was used," the report says. If government agents, hackers or anyone make overly intrusive searches, the evidence will be there for later investigators.

Any automatic audit trail will need a guarantee that it is authentic; one way to do this is to use timestamps. These are electronic signatures that match an electronic file to a point in time. Many timestamps in use are easily forged, but in 1991, Stuart Haber and Scott Stornetta at communications company Bellcore in Morristown, New Jersey, developed a reliable electronic timestamping system. They used mathematical functions called hash functions, which turn a long string of characters such as the bits that encode an electronic file into a shorter fixed-length string, called a hash number, that can stand in for the original string.

To see how hash numbers relate to timestamps, simply leaf through the pages of The New York Times. Surety, a Bellcore spin-off company in Hemdon, Virginia, publishes a timestamp in the commercial section every week. Surety's clients include intellectual property firms and writers' guilds--anyone who expects the time that an electronic file was created to be critical to their business.

Every second, Surety receives hash numbers from its customers; It combines the hash numbers for each second into a superhash value, and then combines that superhash with the superhash for the previous second, and so on. Once a week, the superhash value for that week is published in The New York Times.

The only way to beat the system would be to find a document that had the same hash number as another document. Then someone could claim that this document, not the other, was the one time stamped by Surety. But hash functions are one-way operations that should not have a mathematical inverse. So even if you had the right hash number, you should not be able to work out what the original document said. Surety uses such long hash functions that the chance of guessing a string of bits that would produce a given hash value is 1 in [10.sup.98]-more than the number of atoms in the universe--and computationally impossible.

Given any document, you can convert it into its hash value and look up the second it was registered by the Surety system. There is no way to lie about when a document was created, because the history of the process is publicly documented. "This is still the only [full-]integrity way to do checking of online records," says Haber, who co-founded Surety and is now working for Hewlett-Packard Labs in Princeton, New Jersey. Intelligence agencies could be made to keep similar, tamper-evident logs of their own searches, making them more responsible for their activity.

Of course, even if TIA does create and use all the privacy protection technologies it can, it will not satisfy everyone. David Sobel, a lawyer with the Electronic Privacy Information Center based in Washington DC, says that while there may be technologies that can protect against government spying, TIA will still infringe on people's privacy. TIA is "inherently contrary to the American tradition of criminal justice' he says. To Sobel and many other civil libertarians, it is objectionable for investigators to seek evidence of criminality in private electronic records, whether they find it or not. "The bottom line is the government is making an arbitrary determination that a pattern somehow justifies branding a person a suspect," says Sobel.

This, despite all the best efforts and intentions of the technologists, programmers, wordsmiths and logo designers, is the bottom line. The public is suspicious, alarmed and combative about surveillance technologies because of past experiences. The questionable, and in some cases criminal, government-sanctioned actions of the past suggest that giving government agencies cleverer means of surveillance is guaranteed to lead to abuse.

All of which doesn't leave much hope for those trying to track terrorists. But perhaps if we put as much money and effort into developing and implementing technological restraints for Big Brother as has traditionally gone into equipping him, public attitudes might change. DARPA may not be pleased with Congress's interference, but the agency has been given a golden opportunity to remove the shadow from surveillance technologies. It's not about what you can do. It's what you can't do that counts.