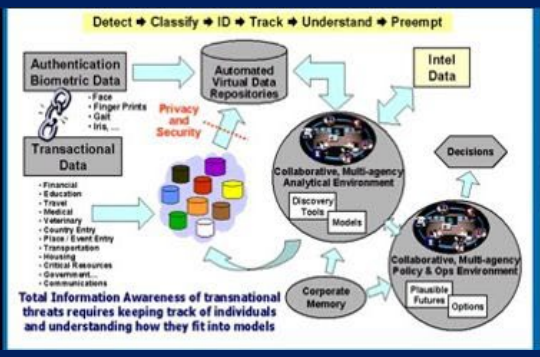
UNCATEGORIZED

The Total Information Awareness Project Lives On

Technology behind the Pentagon's controversial data-mining project has been acquired by NSA, and is probably in use.

By Mark Williams Pontin

April 26, 2006



In April, the Electronic Frontier Foundation (EFF), the advocacy organization for citizens' digital rights, filed evidence to support its class-action lawsuit alleging that telecom giant AT&T gave the National Security Agency (NSA), the ultra-secret U.S. agency that's the world's largest espionage organization, unfettered access to Americans' telephone and Internet communications.

The lawsuit is one more episode in the public controversy that erupted in December 2005, when the *New York Times* revealed that, following September 11, President Bush authorized a far-reaching NSA surveillance program that included warrantless electronic eavesdropping on telephone calls and e-mails of individuals within the United States.

Critics charged that the Bush administration had violated both the Constitution's Fourth Amendment, which protects citizens against unwarranted search or seizure, and the Foreign Intelligence Surveillance Act (FISA) of 1978, which requires eavesdropping warrants to be obtained from a special court of judges empowered for that purpose.

In February 2006, the controversy intensified. Reports emerged that component technologies of the supposedly defunct Total Information Awareness (TIA) project – established in 2002 by the

Pentagon's Defense Advanced Research Projects Agency (DARPA) to develop advanced information technology to counter terrorists, then terminated by Congress in 2003 because of widespread criticism that it would create "Orwellian" mass surveillance – had been acquired by the NSA.

Washington's lawmakers ostensibly killed the TIA project in Section 8131 of the Department of Defense Appropriations Act for fiscal 2004. But legislators wrote a classified annex to that document which preserved funding for TIA's component technologies, if they were transferred to other government agencies, say sources who have seen the document, according to reports first published in *The National Journal*. Congress *did* stipulate that those technologies should only be used for military or foreign intelligence purposes against non-U.S. citizens. Still, while those component projects' names were changed, their funding remained intact, sometimes under the same contracts.

Thus, two principal components of the overall TIA project have migrated to the Advanced Research and Development Activity (ARDA), which is housed somewhere among the 60-odd buildings of "Crypto City," as NSA headquarters in Fort Meade, MD, is nicknamed. One of the TIA components that ARDA acquired, the Information Awareness Prototype System, was the core architecture that would have integrated all the information extraction, analysis, and dissemination tools developed under TIA. According to *The National Journal*, it was renamed "Basketball." The other, Genoa II, used information technologies to help analysts and decision makers anticipate and pre-empt terrorist attacks. It was renamed "Topsail."

Has the NSA been employing those TIA technologies in its surveillance within the United States? And what exactly is the agency doing, anyway?

The hearings that the Senate Judiciary Committee convened in February to consider the NSA's surveillance gave some clues. Attorney General Alberto Gonzales, maintaining the administration's defense against charges that it violated the Fourth Amendment and FISA, told senators, firstly, that Article II of the U.S. Constitution granted a president authority to conduct such monitoring and, secondly, that the Authorization to Use Military Force (AUMF) passed after September 11 specified that the president could "use all necessary and appropriate force" to prevent future terrorist acts. Regarding FISA, Gonzalez claimed, the NSA had sidestepped its requirements to obtain warrants for electronic eavesdropping in particular cases. But, overall, the attorney general said, FISA worked well and the authorities had used it increasingly. The available facts support Gonzalez's contention: while the FISA court issued about 500 warrants per year from

1979 through 1995, in 2004 (the last year for which public records exist) 1,758 warrants were issued.

But when senators asked why, given the fact that FISA had provisions by which government agents could wiretap first and seek warrants later, the Bush administration had sidestepped its requirements at all, Gonzalez claimed he couldn't elaborate for reasons of national security.

Former NASA director General Michael Hayden, in charge when the NSA's surveillance program was initiated in 2002, was slightly more forthcoming. FISA wasn't applicable in certain cases, he told the senators, because the NSA's surveillance relied on what he called a "subtly softer trigger" before full-scale eavesdropping began. Hayden, who is nowadays the nation's second-highest ranking intelligence official, as deputy director of national intelligence, said he could answer further questions only in closed session.

Gonzalez's testimony that the government is making increased use of FISA, together with his refusal to explain why it's inapplicable in some cases – even though retroactive warrants can be issued – implies that the issue isn't simply that government agents may sometimes want to act quickly. FISA rules demand that old-fashioned "probable cause" be shown before the FISA court issues warrants for electronic surveillance of a specific individual. Probable cause would be inapplicable if NSA were engaged in the automated analysis and data mining of telephone and e-mail communications in order to target possible terrorism suspects.
As the Electronic Frontier Foundation's lawsuit against AT&T reveals, NSA has access to the switches and records of most or all of the nation's leading telecommunications companies. These companies' resources are extensive: AT&T's data center in Kansas, for instance, contains electronic records of 1.92 trillion telephone calls over several decades. Moreover, the majority of international telecommunications nowadays no longer travel by satellite, but by undersea fiber-optic cables, so many carriers route international calls through their domestic U.S. switches.

With the telecom companies' compliance, the NSA can today tap into those international communications far more easily than in the past, and in real time (or close to it). With access to much of the world's telecom traffic, the NSA's supercomputers can digitally vacuum up every call placed on a network and apply an arsenal of data-mining tools. Traffic analysis, together with social network theory, can reveal patterns indiscernible to human analysts, possibly suggesting terrorist activity. Content filtering, applying highly sophisticated search algorithms and powerful statistical methods like Bayesian analysis in tandem with machine learning, can search for particular words or language combinations that may indicate terrorist communications.

Whether the specific technologies developed under TIA and acquired by ARDA have actually been used in the NSA's domestic surveillance programs – rather than only for intelligence gathering overseas – has not been proved. Still, descriptions of the two former TIA programs that became Topsail and Basketball mirror descriptions of ARDA and NSA technologies for analyzing vast streams of telephone and e-mail communications. Furthermore, one project manager active in the TIA program before it was terminated has gone on record to the effect that, while TIA was still funded, its researchers communicated regularly and maintained "good coordination" with their ARDA counterparts.

It's this latter fact that is most to the point. Whether or not those specific TIA technologies were deployed for domestic U.S. surveillance, technologies very much like them were. In 2002, for instance, ARDA awarded $64 million in research contracts for a new program called Novel Intelligence from Massive Data. Furthermore, overall, a 2004 survey by the U.S. General Accounting Office, an investigative arm of Congress, found federal agencies operating or developing 199 data mining projects, with more than 120 programs designed to collect and analyze large amounts of personal data on individuals to predict their behavior. Since the accounting office excluded most of the classified projects, the actual numbers would likely have been far higher.

Beyond these programs, additionally, there exist all the data-mining applications currently employed in the private sector for purposes like detecting credit card fraud or predicting health risks for insurance. All the information thus generated goes into databases that, given sufficient government motivation or merely the normal momentum of future history, may sooner or later be accessible to the authorities.

How should data-mining technologies like TIA be regulated in a democracy? It makes little sense to insist on rigid interpretations of FISA. This isn't only because when the law was passed by Congress 30 years ago, terrorist threats on al Qaeda's scale did not yet exist and technological developments hadn't gone so far in potentially giving unprecedented destructive power to small groups and even individuals. Today's changed technological context, additionally, invalidates FISA's basic assumptions.

In an essay published next month in the *New York University Review of Law and Security*, titled "Whispering Wires and Warrantless Wiretaps: Data Mining and Foreign Intelligence Surveillance," K. Taipale, executive director of the Center for Advanced Studies in Science and Technology Policy, points out that in 1978, when FISA was drafted, it made sense to speak exclusively about

intercepting a targeted communication, where there were usually two known ends and a dedicated communication channel that could be wiretapped.

With today's networks, however, data and increasingly voice communications are broken into discrete packets. Intercepting such communications requires that filters be deployed at various communication nodes to scan all passing traffic with the hope of finding and extracting the packets of interest and reassembling them. Thus, even targeting a specific message from a known sender today generally requires scanning and filtering the entire communication flow in which it's embedded. Given that situation, FISA is clearly inadequate because, Taipale argues, were it to be "applied strictly according to its terms prior to any 'electronic surveillance' of foreign communication flows passing through the U.S. or where there is a substantial likelihood of intercepting U.S. persons, then *no automated monitoring of any kind* could occur."

Taipale proposes not that FISA should be discarded, but that it should be modified to allow for the electronic surveillance equivalent of a *Terry* stop – under U.S. law, the brief "stop and frisk" of a person by a law enforcement officer based on the legal standard of reasonable suspicion. In the context of automated data mining, it would mean that if suspicion turned out to be unjustified, after further monitoring, it would be discontinued. If, on the other hand, continued suspicion was reasonable, then it would continue, and at a certain point be escalated so that human agents would be called in to decide whether a suspicious individual's identity should be determined and a FISA warrant issued.

To attempt to maintain FISA and the rest of our current laws about privacy without modifications to address today's changed technological context, Taipale insists, amounts to a kind of absolutism that is ultimately self-defeating. For example, one of the technologies in the original TIA project, the Genisys Privacy Protection program, was intended to enable greater access to data for security reasons while simultaneously protecting individuals' privacy by providing critical data to analysts via anonymized transaction data and by exposing identity only if evidence and appropriate authorization was obtained for further investigation. Ironically, Genisys was the one technology that definitely had its funding terminated and was not continued by another government agency after the public outcry over TIA.

Home page image is available under GNU Free Documentation License 1.2. Caption: Original logo of the now-defunct Total Information Awareness Office, which drew much criticism for its "spooky" images.