I. Introduction

The government holds enormous amounts of information about people's transactions and activities. \(^n1\) "[F]ederal agencies and departments maintain almost 2000 databases, including records pertaining to immigration, bankruptcy, Social Security, military personnel, as well as countless other matters." \(^n2\) The National Directory of New Hires contains information about all people who obtain a new job anywhere in the nation, including Social Security numbers, addresses, and wages. \(^n3\) State governments keep "records of arrests, births, criminal proceedings, marriages, divorces, property ownership, voter registration, and workers compensation," as well as records on \(^n4\) numerous professionals like "doctors, lawyers, engineers, insurance agents, nurses, police, accountants, and teachers."
Private companies also hold tremendous amounts of personal information, some of which comes from government records. \(^5\)

Experian maintains a database of credit information on about 215 million people and demographic information on about 215 million consumers in 110 million U.S. households. \(^6\)

ChoicePoint has 14 billion records on individuals and businesses that can be used for tasks like pre-employment screening of job candidates. \(^7\)

MIB, Inc., has profiles of medical information on about 15 million individuals collected from its association of 600 insurance companies. \(^8\)

In 1996, Catalina Marketing, Inc. began to collect shopping data on an estimated 143 million shoppers per week from more than 11,000 supermarkets nationwide. By 1998, Catalina had reportedly amassed a 2-terabyte database with 18 billion rows of data. \(^9\)

It’s clear that the government wants to use data-mining technologies to analyze this data for law enforcement and anti-terrorism purposes. \(^10\) The ill-fated Total or Terrorism Information Awareness (TIA) program was only the tip of the iceberg. \(^11\)

It’s equally clear that government data mining poses serious threats to civil liberties like privacy, freedom of association, and freedom of speech. \(^12\) The question I address in this article is whether the Fourth Amendment constrains government data mining in any meaningful way. Suppose a law enforcement agent decides to analyze these myriad databases for patterns or clues to potential criminal or terrorist activity. She might begin with an identified suspect and look for evidence of links with other suspects, crimes or suspicious activities. She might hypothesize a pattern that seems to suggest wrongdoing and look for persons whose activity fit that pattern. She might have no suspect or pattern but use the database to look for possible patterns. \(^13\) Does the Fourth Amendment apply to such investigative activities?

It certainly should. After all, the Fourth Amendment embodies procedural values like accountability and control of government discretion as well as substantive values like privacy and free speech precisely because of the Framers’ experience with arbitrary British general warrants and writs of assistance. \(^14\) Given that data mining takes place behind closed doors, it raises transparency problems akin to wiretapping: how can society govern searches that neither targeted individuals nor society at large will know about unless informed by law enforcement?

For data mining, the question is difficult if we assume that the government analyzes lawfully acquired information such as privately compiled databases of financial, communications, and other transactions or government databases of motor vehicle registrations, tax returns, and so on. Under modern Fourth Amendment rules about records held by third parties, we have no reasonable expectation of privacy in any of these records, because we have "knowingly exposed" the information to the public that I made a phone call to my wife at noon last week, or that I wrote a check to buy groceries last month. It seems to follow that government data mining of such databases cannot be a search.

Or does it? I suggest that it does not, that government data mining does implicate the Fourth Amendment. As a threshold matter, even when the Supreme Court has approved of privacy-intrusive government techniques, it has left
open the possibility that overbroad use of such techniques would be treated differently. Equally important, despite the apparent force of the "knowing exposure" doctrine, the mere fact that things or information have been exposed to others does not automatically mean that the Fourth Amendment becomes irrelevant. One's words in a conversation are by definition exposed to the listener, but such private communications receive full Fourth Amendment protection.

Part II of this article briefly describes data-mining operations from a non-technical perspective. Part III then examines the issues at stake in the data mining debate: efficacy; substantive values of privacy, free speech, and freedom of association; and the core procedural value of accountability. Finally, Part IV argues that even if the individual facts in a database are "knowingly exposed to the public," the patterns or inferences uncovered by data mining are not. Data mining is like opening a container to find out what is inside or like scientifically analyzing blood or urine to discover facts about a person, both of which are considered Fourth Amendment searches. It also suggests that we should think about a Fourth Amendment category of "cognition-enhanced" searches, as opposed to sense-enhanced searches.

II. What is data mining?

A recent federal advisory committee report defined government data mining broadly as "searches of one or more electronic databases of information concerning U.S. persons, by or on behalf of an agency or on behalf of an agency or employee of the government." In the technical literature, data mining is more narrowly defined as the "nontrivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data." Similarly, the General Accounting Office defines data mining as "the application of database technology and techniques- such as statistical analysis and modeling-to uncover hidden patterns and subtle relationships in data and to infer rules that allow for the prediction of future results." Both of these definitions seek to distinguish mere information retrieval using traditional query and report tools, which describe what is in a database, from "true" data mining, which uses automated processes to discover patterns. Because such patterns are themselves knowledge, the field is often referred to as "knowledge discovery."

Typically, data mining algorithms are first applied to "training data" in order to find potentially useful patterns. For instance, data mining algorithms can be trained on data for which the correct prediction is already known, like a database containing many cases of both good and bad loans; "given a set of known fraudulent and nonfraudulent credit-card transactions or insurance claims, the computer system may learn general patterns that can be used to flag future cases of possible fraud." For example, an analyst might use a data-mining tool with a loan default database in order to figure out the risk factors for loan default. The analyst might confirm a hypothesis that people with high debt and low incomes were bad credit risks, but the data-mining tool might also discover an unanticipated pattern, such as that age is also a determinant of risk. This unanticipated pattern or model would then be tested on new data in order to test its validity.

The resulting rules or models are then applied to make inferences over data for which the correct prediction is not known, usually called "automated data analysis." Automated data analysis is generally used for either subject- or pattern-based queries. Subject-based analysis resembles traditional investigation: a particular subject is the starting point, and the technology automates the process of finding key relationships or associations. The subject could be a person, a clue to a person like a telephone number or email address, a place or an event.
Pattern-based queries are like the financial examples given above—some model or pattern of behavior is identified and then one searches for instances of that pattern in a database or databases. As with subject-based analysis, the pattern could be based on common sense or past intelligence, and need not be the product of data mining: "Are there foreign visitors to the United States who are staying in urban areas, buying large amounts of fertilizer and renting trucks?" But the major reason for wishing to use data mining is to discover patterns that we humans might not think of on our own; as one observer put it, data mining techniques "can find links, patterns, and anomalies in masses of data that humans could never detect without this assistance." 

The now-defunct TIA program originally included several programs involving both subject- and pattern-based data mining of what TIA called the "transaction space." Undersecretary of Defense Pete Aldridge said that the "[t]he purpose of TIA would be to determine the feasibility of searching vast quantities of data to determine links and patterns indicative of terrorist activities." As Poindexter put it: "If terrorist organizations are going to plan and execute attacks against the United States, their people must engage in transactions and they will leave signatures in this information space." This transaction space included: financial, educational, travel, medical, veterinary, country entry, place/event entry, transportation, housing, critical resources, and communications. Aldridge said that TIA would look for "connections between transactions" (passports, visas, work permits, driver's licenses, credit card, airline tickets, rental cars, gun purchases, chemical purchases) and "events" (arrest or "suspicious activities and so forth").

TIA's Wargaming the Asymmetric Environment (WAE) project, which was aimed at developing models of terrorist behavior, appeared to exemplify "true" data mining: it assumed that different terrorist groups have distinct "styles" or "signatures" that can be identified in the transaction space, and developed indication and warning models for select terrorist groups and individuals based on "their behavior in the broader context of their political, cultural, and ideological environment." Similarly, TIA's Scalable Social Network Analysis (SSNA) project was aimed at developing an algorithm program to help distinguish potential terrorist cells based on their patterns of interactions from legitimate groups of people and identify when a terrorist group plans to execute an attack.

TIA's Evidence Extraction and Link Discovery (EELD) program, on the other hand, focused on "link analysis," a form of subject-based automated data analysis. "EELD is developing detection capabilities to extract relevant data and relationships about people, organizations, locations, and activities from message traffic and open source data. It will then link together related items that comprise potential terrorist groups or scenarios and learn patterns of different groups or scenarios to identify new organizations or emerging threats."

At the DARPATech 2002 conference in Anaheim, the EELD program manager gave a nice example:

If I had access to your transactions, and of other attendees, and perhaps of other travelers to this area for the past few days, or you mine, I could tell if you are traveling alone or with a colleague or with your family. I could tell if you are the type of person who can make advance plans and buy tickets in advance or the type who needs the flexibility to change at the last minute. If I combined registration information I could learn a lot about the demographics of attendance. I could identify DARPA personnel, other Government personnel, and representatives of large and small contractors. I could also find people who fit cleanly into a category, or who don't. With enough information I could perhaps distinguish between people visiting Disneyland for a vacation and those coming to DARPATech.

Other examples of government data-mining programs include:
-the second generation of the Computer-Assisted Passenger Prescreening System (CAPPS II) being developed by
the Department of Homeland Security’s (DHS) Transportation Security Administration (TSA), "which compares airline
passenger names with private- and public-sector databases to assess the level of risk a passenger might pose." n40

-the DHS's mandate to "establish and utilize . . . data-mining and other advanced analytical tools" to detect and
identify threats of terrorism against the United States. n41

-the Multistate Anti-Terrorism Information Exchange (MATRIX) program begun by the state of Florida in
conjunction with data-aggregation company Seisint Inc., which links law enforcement records with other government
and private-sector databases, and is designed to "find patterns and links among people and events faster than ever
before." n42

[*398]

III. The constitutional values at stake in the data mining debate

It should be obvious from the foregoing that data mining to discover behavioral patterns raises serious
constitutional concerns for the individuals whose activities are recorded in the database. After all, "the very definition of
a 'pattern of behavior' is one in which data associated with a given individual are grouped so that a trip to Location A
can be seen in the context of Transaction B and a phone call with Person C. It is exactly such linkage that gives rise to
privacy concerns." n43

Moreover, data mining does not only challenge substantive values like privacy; it also challenges procedural values
like particularized suspicion. As the TAPAC Report noted, "[t]he common feature of all of these programs is that they
involve sifting through data about identifiable individuals, even though those individuals have done nothing to warrant
government suspicion, in search of useful information." n44

A. Privacy and associational freedom

Scholars have explained how the market for personal information has led to the creation of extensive "digital
dossiers" that are increasingly used by the government for law enforcement and other purposes. n45 An obvious harm
associated with unlimited government access to information is the threat of undue social control (the "Big Brother"
metaphor). n46

Of particular concern here is the way that digital dossiers reveal much about our lives and personal associations.
"Detailed records of an individual's reading materials, purchases, diseases, and website activity enable the government
to assemble a profile of an individual's finances, health, psychology, [ *399] beliefs, politics, interests, and lifestyle.
This data can unveil a person's anonymous speech and personal associations." n47

The courts have long recognized, of course, that there is a "vital relationship between freedom to associate and
privacy in one's associations," n48 This concern is especially great in the counter-terrorism context, because "[n]ational
security cases . . . often reflect a convergence of First and Fourth Amendment values not present in cases of 'ordinary'
crime. Though the investigative duty of the executive may be stronger in such cases, so also is there greater jeopardy to
constitutionally protected speech." n49
The risk to associational privacy is obvious when one considers the kind of information in our digital dossiers. For instance, Acxiom claims that its InfoBase profiler product contains the following personal information: name; address; phone number; occupation; date of birth; latitudinal/longitudinal coordinates; gender; ethnicity; age; income; net worth; political party; height and weight; education; marital status; whether one wears corrective lenses; whether subject rents or owns dwelling; years of residence; value of home; mortgage amount and interest rate; home loan to value ratio; date home was built and date purchased; square footage of home and lot; whether home is located in a census tract where more than 50 percent of households are non-white; adult age ranges in household; number of adults in household; children's age ranges in household; number of children in household; number of generations in household; total number of occupants in household; whether there is a "working woman" in household; which credit cards subject owns; range of new credit; where subject likes to shop; model and make of automobile (including a "lifestyle indicator" designation based on the type of car); blue book value of vehicle; whether subject has a history of buying new cars; whether subject buys items through mail order and in what dollar amounts; whether subject owns a cat or a dog; whether subject donates to charities; whether subject owns real estate investments; whether subject has stock/bond investments; whether subject is a military veteran; whether subject likes to gamble, sew, garden, watch television, hunt, jog, sail, diet, play video games, drink wine, or read the Bible; and whether subject's overall "lifestyle composite" classifies him/her as a "Traditionalist," "Connoisseur," "Chiphead" (like computers and science), or member of the "Intelligentsia." n50

Moreover, data mining can easily circumvent existing privacy protections. A recent state case imposing a warrant requirement on police use of GPS tracking devices noted that:

The device can provide a detailed record of travel to doctors' offices, banks, gambling casinos, tanning salons, places of worship, political party meetings, bars, grocery stores, exercise gyms, places where children are dropped off for school, play, or day care, the upper scale restaurant and the fast food restaurant, the strip club, the opera, the baseball game, the 'wrong' side of town, the family planning clinic, the labor rally. In this age, vehicles are used to take people to a vast number of places that can reveal preferences, alignments, associations, personal ails and foibles. The GPS tracking devices record all of these travels, and thus can provide a detailed picture of one's life. n51

The Washington court was right, but it will often be unnecessary for the government to track us, because for most of us much of our lives are already described in transactional databases. n52

B. Accountability: Data mining as covert surveillance

The Fourth Amendment's "basic purpose . . . is to safeguard the privacy and security of individuals against arbitrary invasions by governmental officials." n53 To control official discretion, the Fourth Amendment relies on accountability mechanisms that are fundamentally concerned with the need for an objective justification of government behavior. The warrant requirement is the most obvious example: the traditional search warrant issues only after a neutral, detached magistrate finds that the government has made an adequate factual showing of probable cause. n54 In addition, the warrant itself must define the scope of the search with particularity, thus ensuring that the search will be tailored to the magistrate's probable-cause determination. n55

The Supreme Court recently reiterated the importance of particularity to the warrant process in finding unconstitutional a warrant that did not correctly describe the items to be seized, even though the supporting affidavits contained particular descriptions. As the Court put it, "unless the particular items described in the affidavit are also set forth in the warrant itself (or at least incorporated by reference, and the affidavit present at the search), there can be no written assurance that the Magistrate actually found probable cause to search for, and to seize, every item mentioned in
The Fourth Amendment search warrant procedure can be viewed as a more general accountability model with several key default parameters. First and foremost is the need for particularized suspicion: there must be facts that demonstrate a good reason to search this person, place, or thing. Second, this factual justification must meet some standard of certainty or likelihood, e.g. "probable cause." Third, the warrant itself must describe with particularity the scope of the search. Fourth, there must be an independent check, e.g., the requirement of a neutral and detached magistrate, which ensures the objectivity of the determinations that the justification exists, that it meets the requisite certainty standard, and that the scope of the search is objectively defined. Fifth, this independent assessment should take place before the government conducts its search.

1. Objective justification and the particularity problem

Particularity is crucial to Fourth Amendment accountability because it "makes general searches . . . impossible and prevents the seizure of one thing under a warrant describing another. As to what is to be taken, nothing is left to the discretion of the officer executing the warrant." It is well settled that this aspect of Fourth Amendment justification is a direct response to the British "general warrants" and "writs of assistance," which could be issued without any suspicion of illegal activity and were "unlimited geographically and perpetual temporally."

The Fourth Amendment also generally insists on some degree of particularized suspicion. Here again, control of official discretion is the primary concern. Lack of particularity frustrates "the purpose of the probable-cause requirement of the Fourth Amendment, to keep the state out of constitutionally protected areas until it has reason to believe that a specific crime has been or is being committed . . . ." The secondary concern is the quality of official justification, to assure that mistakes are "those of reasonable men, acting on facts leading sensibly to their conclusions of probability."

The Court's treatment of a state eavesdropping statute in Berger v. New York is illustrative. The statute required prior approval by a judge and could have been construed to require probable cause, but the Court focused on its lack of any particularity safeguards and found that the statute "actually permits general searches by electronic devices . . . ." More generally, the Court observed that "[t]he need for particularity and evidence of reliability in the showing required when judicial authorization of a search is sought is especially great in the case of eavesdropping," which "[b]y its very nature . . . involves an intrusion on privacy that is broad in scope."

The use of patterns discovered through data mining raises similar particularity issues. Imagine a database of a million people and a hypothesis that those who meet certain criteria are highly likely to be terrorists. But you don't know whether any of these million people actually do meet these criteria; if you did, you wouldn't need to run the search. The basic problem is lack of particularized suspicion: data about these persons would be "searched" without any reason to believe either that the database contains evidence of terrorist activity or that any person "in" the database is a terrorist. Like eavesdropping, pattern-oriented data mining (or automated data analysis) by its very nature involves broad intrusions on privacy, and demands careful attention to particularity.

Even when automated data analysis is subject-oriented-as when the government is investigating a particular suspect or incident-particularized suspicion remains a problem. If the government has reason to believe that "John Smith" is a terrorist, it has particularized suspicion as to him. If the government reasonably believes that someone who uses a particular phone number or email address is a terrorist, again there is some particularized suspicion.
But how far does that particularized suspicion get you? Link analysis, for instance, focuses on the transactional connections between the subject and other people: who lives with John Smith, who corresponds with johnsmith@aol.com, and so on. Does the mere fact that Jane Doe has certain connections to John Smith mean that there is particularized suspicion as to her?

Under current law, that seems highly unlikely. In Ybarra v. Illinois, the police had a valid search warrant to inspect a tavern and a specific bartender based on probable cause to believe that the bartender would have heroin for sale. Upon entering the tavern, the officers announced their purpose and advised those present that they were going to conduct a cursory search for weapons. An officer frisked each of the customers present in the tavern. The frisk of Ybarra revealed "a cigarette pack with objects in it," which turned out to be foil packets of heroin. The Court held the search of Ybarra to be illegal given that "the agents knew nothing in particular about Ybarra, except that he was present" in the tavern.

[A] person's mere propinquity to others independently suspected of criminal activity does not, without more, give rise to probable cause to search that person. Where the standard is probable cause, a search or seizure of a person must be supported by probable cause particularized with respect to that person. This requirement cannot be undercut or avoided by simply pointing to the fact that coincidentally there exists probable cause to search or seize another or to search the premises where the person may happen to be.

The Ybarra principle would clearly frustrate simple link analysis that begins with a person for whom particularized suspicion exists and then follows links to other persons who merely have associated with the original suspect. Accordingly, the TAPAC Report recognized that "the power of data mining technology and the range of data to which the government has access have contributed to blurring the line between the subject- and pattern-based searches . . . [e]ven when a subject-based search starts with a known suspect, it can be transformed into a pattern-based search as investigators target individuals for investigation solely because of their connection with the suspect."

The particularity problem doesn't end with particularized suspicion, of course. Even with particularized suspicion, the search itself must have a particularized scope; the searcher's discretion must be limited and defined. Under the warrant requirement, the places or persons to be searched must be particularly described so that the magistrate can determine on the record that there is an objective justification to search those places or persons. This requirement does not make automated data analysis impossible, but it does require considerable safeguards.

Again, Berger is instructive. The Court in Berger compared New York's procedure to that approved in Osborn v. United States, where the court order "described the type of conversation sought with particularity," so that "the officer could not search unauthorized areas," and "once the property sought . . . was found the officer could not use the order as a passkey to further search." Moreover, "the order authorized one limited intrusion rather than a series or a continuous surveillance," and "a new order was issued when the officer sought to resume the search and probable cause was shown for the succeeding one." Finally the officer was required to and did make a return on the order showing how it was executed and what was seized. These sorts of precautions would seem absolutely necessary to ensure that "the search as actually conducted 'was reasonably related in scope to the circumstances which justified the interference in the first place.'"

One data mining proponent argues that, from a particularity perspective, automated data analysis is like a police officer's noticing that a masked man is running on a public street; if stopping and questioning the masked man is reasonable, why isn't it reasonable to investigate someone whose electronic trail indicates suspicion? But this example misses the point, even if we assume that automated data analysis involves matters in "plain view." When the
officer observes the running masked man, the masked man's behavior is the particular trigger for investigating him and not someone else. An automated database "search" is not based on particularized suspicion; it is based on the hope that suspicious connections or patterns will be found. The analysis may yield particularized suspicion, but the whole point is that you suspect before you search.

The inherent particularity problem is that the government is running queries, whether subject- or pattern-oriented, over a database or set of databases containing personal information about many people who are not suspected of anything. Indeed, DARPA eventually admitted that the data mining originally contemplated by TIA "must inevitably lead to 'fishing expeditions' through massive amounts of personal data and a wholesale invasion of Americans' privacy that yields, basically, nothing in terms of finding terrorists." n82

2. Transparency and visibility

Traditionally, the act of and consequences of searching created their own public accountability. Knowledge of searches is relatively public when government physically searches homes or persons; physical invasions are hard to hide. Moreover, searches have traditionally been used to combat ordinary crime, meaning that prior judicial review is often augmented by after-the-fact [*406] judicial review on motions to suppress. "The idea is much the same as the one behind requiring public trials: Visibility is a powerful regulatory tool." n83

But modern "search" activity is far less visible to us; n84 electronic surveillance easily operates without a target's knowledge. We will only know about wiretapping-or data mining and automated data analysis-if the government decides to tell us about it.

When the ultimate goal of a search is to go to trial, there will be some accountability, even though post hoc review is a poor substitute for prior review. n85 But even the possibility of truly objective post hoc review cannot provide accountability when the fruits of a search are not used in court. n86

More disturbingly, accountability fails when the police dissemble or otherwise abuse their power. As the Court has said, "The Fourth Amendment does not contemplate the executive officers of Government as neutral and disinterested magistrates. . . . The historical judgment, which the Fourth Amendment accepts, is that unreviewed executive discretion may yield too readily to pressures to obtain incriminating evidence and overlook potential invasions of privacy and protected speech." n87

A modern case in point is Whitaker v. Garcetti, n88 which illustrates both the broad sweep of electronic surveillance and the potential for abuse. In Garcetti, L.A. police units used two judicially authorized wiretap orders, one that intercepted 30,000 conversations over 11 months, and another that intercepted "dozens of thousands of conversations" over 22 months, to gather evidence of imminent criminal conduct about persons who were not named in the wiretap orders. n89 Instead of acting on this evidence, however, they [*407] "handed off" the information to a second police unit without saying that the information had been obtained via a wiretap. The second unit would then either make an arrest or obtain a search warrant without revealing the existence of the wiretap. n90

Defendants in 58 cases were never informed that their prosecutions were based on wiretaps until the issuance of a state court order after an in camera proceeding outside the presence of defense counsel. n91 As the court put it, "[t]he [criminally] accused is . . . never informed of or able to challenge the affidavit, the wiretap order, or the wiretap, itself (notwithstanding the fact that these are the investigative mechanisms out of which his prosecution originally arose)." n92
The district court found this "hand off" procedure unconstitutional, because "a criminal defendant ha[s] a constitutional right to know that he has been subjected to a Fourth Amendment search from which the investigation against him originally arose[.]" n93

The very technology of data mining and automated data analysis arguably poses even greater challenges for transparency. Profiling algorithms used to rank persons as "threats" or discover "signals" of terrorist activity can be as invisible as wiretaps if care is not taken to maintain faithful records of the operations. And if neural networks are used to "learn" and identify behavior patterns, n94 those triggering patterns will be even more opaque. The advantage of neural networks is that "they offer a means of efficiently modeling large and complex problems in which there may be hundreds of predictor variables that have many interactions." n95 Unfortunately, because there are so many variables, "the parameters become uninterpretable and the network serves as a 'black box' predictor. . . . There is no explicit rationale given for the decisions or predictions a neural network makes." n96

The general problem can be stated as follows: Assume that automated data analysis casts suspicion on someone, causing the person harm. Suppose the person challenges the government action. Presumably, the government must establish the reasonableness of the data mining, as it would need to explain how a forensic laboratory concluded that a crime-scene fingerprint [*408] identified a suspect. How does the government show how the data mining operation works and that it reasonably cast suspicion on the person?

Indeed, if the patterns, models or processes used to decide that a person is suspicious are considered secret, it may be impossible to determine precisely how an individual has been singled out for investigation. The MATRIX program, for instance, apparently used a "high terrorism factor" scoring system that identified 120,000 people as being statistically likely to be terrorists and led to some investigations and arrests, but administrators now say that this scoring system is no longer being used. n97 It's unclear that we'll ever know how this system actually worked.

IV. Data mining as a search

The above discussion, I hope, strongly suggests that data mining raises serious Fourth Amendment concerns. But under existing law the Fourth Amendment simply won't apply to data mining unless it can be considered a "search."

I argue that data mining is a search, even when the government has lawfully acquired the individual facts in the database being mined, because the patterns or inferences discovered via data mining often deserve to be private n98 and go beyond the information that can fairly be said to be "knowingly exposed" to others. This claim may sound preposterous. Isn't all the information in the database already completely exposed?

Yes and no. As noted earlier, the data-mining literature distinguishes between traditional "query and report tools," which describe what is in a database, and "true" data mining, which identifies "valid, novel, potentially useful and ultimately understandable patterns in data." n99 Because "such patterns are themselves knowledge," it is completely reasonable to say that finding such patterns or relationships exposes information that is not "in" the database. After all, since 9/11 we have been engaged in a public debate about needing to do a better job of "connecting the dots." Merely having or knowing individual facts is one thing; discerning patterns or relationships within or among facts is another. n100

[*409]

Consider, for instance, a database of purchases created by aggregating bank and credit-card transactional records.
Inspecting this database reveals the following patterns: Jane Doe had bought birth control pills every 3 months for years; during this period, she began buying bridal magazines and then changed her name and established a joint checking account with someone who is apparently her husband; at some point, she stopped buying birth control pills; after about a year, she began buying pregnancy test kits; and then at some point she stopped buying pregnancy test kits and began buying baby clothes and other things at "Babies 'R' Us."

The story that these patterns reveal is obvious: she got married, intended to have a child, and is now expecting. Other inferences are possible, but considerably less plausible. For my purposes, the question is whether she can be said to have "knowingly exposed" the "fact" that she is pregnant. It's not as though the transactional records that make up the database include the results of her home pregnancy test or her doctor's notes confirming her pregnancy. The single most probative transaction in the database—that she bought a home pregnancy test kit—is perfectly consistent with the alternative hypothesis that she was simply buying it for a friend or relative.

My point is simple: because the fact "Jane Doe is pregnant" is not actually in the database, we must question whether she knowingly exposed the fact of her pregnancy to anyone purely by virtue of her having made these purchases. And if she did not "knowingly expose" the fact of her pregnancy, then this fact should remain the subject of Fourth Amendment protection.

More generally, it is not hard to imagine a database of transactions for which it can be said that patterns or potential facts (inferences) are not apparent until one tries to "connect the dots," whether for a targeted individual or among many. Equally generally, these patterns or inferences revealed by analyzing the data in databases are not knowingly exposed to the public. In data mining, "[t]he individual records are almost meaningless in themselves. Instead, only the network of relationships among the people, places, things, and events forms a meaningful pattern." n101

In short, the proper response to the obstacle posed by the "knowing exposure" doctrine is that the underlying patterns or associations are not the same as the surface facts in the database, and that these patterns may themselves be "private."

Perhaps surprisingly, this reasoning is supported by a number of Supreme Court cases. The general threshold point is that when the government attempts to obtain evidence, the Fourth Amendment can be relevant at several levels. n102 Indeed, even when a police officer was lawfully present in a [*410] suspect's apartment "littered with drug paraphernalia" and where "[a] .25-caliber automatic pistol lay in plain view on the living room floor," the Supreme Court found that moving a turntable to read the serial number was a separate search for which probable cause was needed. n103

As noted earlier, individuals generally retain a privacy expectation in the contents of closed containers, and the mere fact that the government has lawful possession of the container does not automatically permit the government to open the container without a search warrant. n104

Computers and other electronic devices have been treated as information containers. n105 Some courts have held that each computer file is, in essence, its own container: "Because computers can hold so much information touching on many different areas of a person's life, there is a greater potential for the 'intermingling' of documents and a consequent invasion of privacy when police execute a search for evidence on a computer." n106 Lawful possession of the computer does not, on this view, automatically permit search of all the files, even if the files are not encrypted or password-protected.
A particularly suggestive case is Walter v. United States, where a shipment of securely sealed packages containing allegedly obscene 8-mm films was mistakenly delivered to a third party rather than the consignee. Employees of the third party opened each of the packages and found individual boxes containing the films; one employee opened a few of the boxes and "attempted without success to view portions of the film by holding it up to the light." The films were turned over to the FBI, which then viewed the films on a projector.

Even though the private employees had already opened the film boxes and unsuccessfully tried to view the films, and even though the FBI had lawful possession of the films, the Court held that the FBI's warrantless viewing of the films was a search that violated the Fourth Amendment. "The private search merely frustrated [the consignor's] expectation in part. It did not simply strip the remaining unfrustrated portion of that expectation of all Fourth Amendment protection. Since the additional search conducted by the FBI—the screening of the films—was not supported by any justification, it violated that Amendment." n109

Another good example is Bond v. United States, where a Border Patrol agent boarded a bus to check the immigration status of the passengers and squeezed the soft luggage in the overhead storage space. The agent detected a "brick-like" object in Bond's canvas bag that turned out to be a "brick" of methamphetamine. On review of Bond's conviction, the Court found that while a bus passenger "clearly expects that his bag may be handled," "[h]e does not expect that other bus passengers or bus employees will, as a matter of course, feel the bag in an exploratory manner." n111 Thus, the agent's "probing tactile examination" of the luggage was an unlawful search.

The analogy to automated data analysis is not perfect, of course. But Walter, Bond, and the container cases clearly make it impossible to say that it cannot be a search to discover "hidden" information "inside" something legitimately possessed. Any impediment to perception, from deliberately placed containers to the mere physical placement of a turntable to the need for specialized equipment in order to extract information, can in the right circumstances transform government acquisition of information into a search.

In non-search cases, the Court has recognized that the aggregation of data can create legally cognizable privacy threats. In a Freedom of Information Act case, the Court found that there was an individual privacy interest in criminal history information summarized in FBI "rap sheets" even when the withheld information was publicly available at the source:

In an organized society, there are few facts that are not at one time or another divulged to another. . . . Plainly there is a vast difference between the public records that might be found after a diligent search of courthouse files, county archives, and local police stations throughout the country and a computerized summary located in a single clearinghouse of information.

The most pertinent line of cases, however, involves scientific testing or analysis of substances. In the drug-testing cases, the Court has found that urine testing involves two separate searches: the initial collection of the urine, and the subsequent chemical analysis of the urine. The urinalysis is a search because "chemical analysis of urine, like that of blood, can reveal a host of private medical facts about an employee, including whether she is epileptic, pregnant, or diabetic." n120

By contrast, the chemical analysis that identified an unknown powder as cocaine in United States v. Jacobsen was not a search. The crucial point for the Court was that the test "could disclose only one fact previously unknown to the agent—whether or not a suspicious white powder was cocaine. It could tell him nothing more, not even
whether the substance was sugar or talcum powder." n122 Thus, the test "compromises no legitimate privacy interest" because a positive finding would only reveal that the substance was cocaine- contraband in which no person can have a legitimate Fourth Amendment interest-while a negative finding would "merely disclos[e] that the substance is something other than cocaine," but "no other arguably 'private' fact . . . ." n123

The obvious conclusion is that whether scientific testing is a search turns upon the technique's potential to discover otherwise imperceptible private facts. And from this perspective, data mining is a lot like urinalysis. Data mining that discovers novel patterns discovers knowledge that is not, strictly speaking, "in" the database. Or that knowledge is "in" the database in the same way that private medical facts about a person's being epileptic, pregnant or diabetic are "in" her urine. They are not on the surface or exposed to view; more is required to get at them.

This reasoning suggests a more general point about the Fourth Amendment treatment of new information technologies that raise privacy concerns. We typically focus on how technology enhances our senses to enable more efficient or extensive collection of information, either by amplifying the range or scope of our natural senses or creating new, somewhat analogous senses. n124

But sensation or information collection isn't everything when it comes to privacy invasion. We should think of these kinds of privacy invasions, as well as scientific or device-enabled analysis-whether of bodily fluids or of databases-as "cognition-enhanced" searches that deserve their own Fourth Amendment jurisprudence. n125

[*414]

The aggregation and organization of already collected information is an obvious cognition enhancement. Fifth Amendment cases like Whalen v. Roe n126 and FOIA privacy exemption cases like United States Dep't of Justice v. Reporters Comm. for Freedom of the Press n127 recognize a privacy interest in the practical obscurity that we have reasonably come to expect in scattered bits of information. n128

The use of recording devices is another kind of obvious cognition enhancement. Making recordings of people's activities concretely memorializes, in an enduring and more transmissible way, conversations or events that would normally persist only in a person's memory. Interestingly, the federal statute that governs surveillance of oral conversations n129 speaks in terms of an expectation of "noninterception," not an expectation of privacy, thus distinguishing between mere overhearing and the use of a device to hear. n130 A person can lack a valid privacy expectation in his normal conversations because others are nearby but still have a valid expectation of noninterception. n131

The concern, of course, is to distinguish between traditional police work that looks for connections from automated data analysis that looks for connections. Data mining proponents argue that many ordinary police investigations involve combing through records to discover connections that indicate illegal activity. Should the use of computers be treated differently?

[*415]

There's good reason to think so. The data only exists in this quantity and quality and can only be analyzed this way because of computers. Data mining and automated data analysis cannot be done without computers. Cheap surveillance "simply makes it too easy, without the loss of a lot of shoe leather and the other costs police traditionally have had to take into account in determining the realistic limits upon their enforcement activities, to engage in random and wholesale snooping. Posnerian cost-benefit balancing is thus no longer a sufficient deterrent to such enlarged investigative strategies, and this is precisely why this activity needs to be brought within the purview of the [F]ourth [A]mendment."
Moreover, if suspicion is not required we will lack the power to hold government accountable for automated data analysis, given its visibility problems. In the modern age of information technology, we need Fourth Amendment doctrines that address information privacy concerns beyond the mere collection of information.

V. Conclusion

Data mining is all the rage today, but it presents serious challenges to both substantive and procedural constitutional values. Legislation may address these issues, but it will probably be insufficient without a constitutional foundation that emphasizes the importance of particularity and public accountability. This essay suggests how one might conceive of a Fourth Amendment basis for regulating data mining as a search.

Legal Topics:

For related research and practice materials, see the following legal topics:
Criminal Law & ProcedureSearch & SeizureSearch WarrantsIssuance by Neutral & Detached MagistratesEducation LawFaculty & StaffCompensationPaymentEnergy & Utilities LawMining IndustryGeneral Overview

FOOTNOTES:


n3 Sweeney, supra note 1, at 14-15 (noting that while the goal of the new hires database was "to better track down parents who owe child support," it holds "information on almost all working Americans, the vast majority of whom have been accused of nothing.").

n4 Solove, supra note 2, at 1403.

n5 "Name and address information come from voting records; land titles are a source of home ownership information; property taxes can
give you assessed values of homes; birth and death records give you information about an individual's parents. The list goes on, there are occupational license records, motor vehicle records that can tell you about an individual's make and model of an automobile, voter registration gives you party political affiliation, and hunting and fishing licenses, boat and airplane licenses can give you information about how a person likes to spend their leisure time. There may be considerably more information available in public records about an individual who has interacted with the courts as a criminal defendant, as a plaintiff or defendant in civil litigation, in a divorce proceeding, as a juror, as the beneficiary of a will.” Transcript of Federal Trade Commission Public Workshop, The Information Marketplace: Merging and Exchanging Consumer Data, 58-59 (Mar. 13, 2001) (statement of Paula Bruening, staff counsel, Center for Democracy and Technology), at http://www.ftc.gov/bcp/workshops/info-mktplace/transcript.pdf (last visited July 3, 2004).


n9 Sweeney, supra note 1, at 9.

n10 The FBI buys files of information about individuals from ChoicePoint, Inc., a major data aggregation company that "cull[s], sort[s] and packag[es] data on individuals from scores of sources, including credit bureaus, marketers and regulatory agencies," and "FBI agents also can go to a dedicated Web intranet site for help in conducting their own searches." Markle Foundation, Protecting America's Freedom in the Information Age 123 (Oct. 2002) (quoting Glenn R. Simpson, Big Brother-in-Law: If the FBI Hopes to Get the Goods on You, It May Ask ChoicePoint- U.S. Agencies' Growing Use Of Outside Data Suppliers Raises Privacy Concerns, Wall Street Journal (Apr. 13, 2001)), at http://www.markletaskforce.org/documents/Markle Full Report.pdf (last visited July 6, 2004). In the counter-terrorism context, the FBI is particularly interested in private-sector data from: the travel industry (airlines, rail, car rental); the telecommunications industry (cellular, land line, Internet); the financial industry (banks, credit card, money transmitters, casinos, brokerage firms); and the services industry (insurance, pharmaceuticals, weapons, chemicals, precursors). Id. at 123-124 (FBI interview with Markle Task Force and Abt Associates).


n12 Strictly speaking, “data mining” refers to the analysis of data for useful patterns, while “automated data analysis” refers to the process of applying these patterns to new data. For more discussion, see text at note 20-27, infra.


n16 For a thorough discussion of this doctrine, see Sherry F. Colb, What is a Search? Two Conceptual Flaws in Fourth Amendment Doctrine and Some Hints of a Remedy, 55 Stan. L. Rev. 119 (2002); id. at 122 (“the Court denies privacy in whatever people 'knowingly expose' to the public. If a person knowingly exposes some object or activity to the public, there has accordingly been no search.”) (footnotes omitted); Wayne R. LaFave, The Forgotten Motto of Obsta Principiis in Fourth Amendment Jurisprudence, 28 Ariz. L. Rev. 291, 301-304 (1986) (criticizing “knowing exposure” doctrine); id. at 304 (“there is a dramatic difference, in privacy terms, between revealing bits and pieces of information sporadically to a small and often select group for a limited purpose and a focused police examination of the totality of that information regarding a particular individual.”).

n17 See, e.g., California v. Ciraolo, 476 U.S. 207, 215 n.3 (1986) ( naked-eye aerial surveillance of fenced-in backyard was not a search, but expressing concern for technology that "discloses . . . intimate associations, objects or activities") (citation omitted); Dow Chemical Co. v. United States, 476 U.S. 227, 238 n.5 (1986) (aerial surveillance of corporate facilities using precise cameras was not a search, but pictures of faces or class rings would raise "more serious privacy concerns"); id. at 239 (device that heard and recorded confidential discussions would raise "far more serious questions"); United States v. Knotts, 460 U.S. 276, 284 (1983) (holding that police use of beepers to track car's movements on public streets was not a search but noting that "if such dragnet-type law enforcement practices as respondent envisions should eventually occur, there will be time enough then to determine whether different constitutional principles may be applicable."); United States v. Miller, 425 U.S. 435, 444 n.6 (1976) (holding that depositor's bank records were not protected by the Fourth Amendment, but noting that the court was "not confronted with a situation in which the Government, through 'unreviewed executive discretion,' has made a wide-ranging inquiry that unnecessarily '[touch]es upon intimate areas of an individual's personal affairs.' Here the Government has exercised its powers through narrowly directed subpoenas duces tecum subject to the legal restraints attendant to such process." (quoting California Bankers Ass'n. v. Shultz, 416 U.S. 21, 78-79 (1974) (Powell, J., concurring))).


n19 TAPAC Report, supra note 11, at 4 (footnote omitted).


n22 An obvious example is the use of simple search queries in Lexis or Westlaw legal databases to find relevant cases and statutes. See Two Crows Corp., Introduction to Data Mining and Knowledge Discovery 3 (1999) [hereinafter Two Crows], at http://www.twocrows.com/intro-dm.pdf (last visited July 31, 2004).

n23 Some algorithms generate probabilistic rules ("a person who makes at least five bank deposits of more than $ 5000 in one day for someone else is 50% likely to be a smurf"); other algorithms build models represented as equations, classification trees, or graphical networks of associations. See National Academy of Sciences, Information Technology for Counterterrorism: Immediate Actions and Future Possibilities 68 (2003) [hereinafter NAS Report], at http://www.nap.edu/html/IT counterterror/ (last visited Jul. 6, 2004) ("Decision-tree learning, neural-network learning, Bayesian-network learning, and logistic-regression-and-support vector machines are among the most widely used statistical machine-learning algorithms.").

n24 Id. Data mining can also be used to predict what individuals are most likely to make certain purchases by analyzing other individuals' past purchasing data. Id.; see Zarsky, supra note 20 (noting how online bookseller Amazon.com uses "association rules" to predict a person's tastes in movies and music by "mining" all the shopping carts used at the website and observing which products tend to be purchased together or by similar people). For a simple illustration of the kinds of models and rules that a money-laundering detection system might use, see David Jensen, Data Mining in Networks, slides 5-9, at http://kdl.cs.umass.edu/people/Jensen/papers/nrcdbsse02.html (last visited July 4, 2004).

n25 Two Crows, supra note 22, at 3.


n27 Much of the explanation in this paragraph is taken from Jensen, supra note 24, at slide 9, at http://kdl.cs.umass.edu/people/jensen/papers/nrcdbsse02/slide09.html (last visited July 4, 2004).

n28 Tether testimony supra note 14.
n29 DeRosa, supra note 26, at 6. Data mining for counterterrorism raises special problems as compared to commercial data mining. First, the data will often be "relational" rather than "propositional," that is, relationships among people, places, things, and events, rather than facts about a given person. Jensen, supra note 24, at slide 21. Second, while commercial data mining typically uses data about individuals to make inferences about individuals, data mining for counterterrorism often seeks to identify high-level things (networks or organizations) based on low-level data about people, places, things, and events. Id. at slide 21-22.

n30 After TIA became a cause celebre, DARPA Director Anthony Tether seemed to retreat from that vision, acknowledging that data mining is "ill-suited" to the task of "find[ing] extremely rare instances of patterns across an extremely wide variety of activities-and hidden relationships among individuals." Tether testimony, supra note 14. Whether TIA's data-mining projects (as opposed to TIA as a whole) are actually defunct is questionable at the time of this writing. See Shannon R. Anderson, Bill of Rights Defense Committee, Total Information Awareness and Beyond: The Dangers of Using Data Mining Technology to Prevent Terrorism, (2004), at http://www.bordc.org/data-mining.pdf (last visited July 18, 2004).

n31 In the counter-terrorism context, pattern-based data mining "involve[s] developing models of what terrorist behavior might look like and then examining databases for similar patterns." TAPAC Report, supra note 11, at 45.


n34 Id.

n35 Aldridge, supra note 32.


n38 DARPA Fact File, supra note 36.


n42 Id. (citing Robert O'Harrow, Jr., U.S. Backs Florida's New Counterterrorism Database, Washington Post, Aug. 6, 2003, at A1); see GAO Report, supra note 21, at 5 ("Information in MATRIX databases includes criminal history records, driver's license data, vehicle registration records, incarceration records, and digitized photographs. Public awareness of MATRIX and of similar large-scale data mining or data mining-like projects has led to concerns about the government's use of data mining to conduct a mass 'dataveillance'-a surveillance of large groups of people-to sift through vast amounts of personally identifying data to find individuals who might fit a terrorist profile.") (footnote omitted).

n43 NAS Report, supra note 23, at 74; see GAO Report, supra note 21, at 6 ("Mining government and private databases containing personal information creates a range of privacy concerns. Through data mining, agencies can quickly and efficiently obtain information on individuals or groups by exploiting large databases containing personal information aggregated from public and private records.").

n44 TAPAC Report, supra note 11, at 4.


n46 See generally Solove, supra note 2 (suggesting that allusion to Kafka's The Trial captures concerns about unaccountable bureaucratic discretion not covered by Orwell's 1984); Peter P. Swire, Financial Privacy and the Theory of High- Tech Government Surveillance, 77 Wash. U. L.Q. 461, 471 (1999) (articulating problems of "how an authoritarian or totalitarian government might use and abuse information about citizens' financial transactions").
n47 Solove, supra note 45, at 1084 (footnotes omitted).

n48 NAACP v. Alabama, 357 U.S. 449, 462 (1958) ("Inviolability of privacy in group association may in many circumstances be indispensable to preservation of freedom of association, particularly where a group espouses dissident beliefs.").


n50 Andrew J. McClurg, A Thousand Words are Worth a Picture: A Privacy Tort Response to Consumer Data Profiling, 98 Nw. U. L. Rev. 63, 76 (2003) (footnote omitted); id. at 77 (noting that Acxiom also possesses unlisted phone numbers, Social Security numbers, email addresses, and magazine subscription lists) (footnotes omitted).


n52 For a more extensive discussion of the civil liberties risks associated with data mining, see TAPAC Report, supra note 11, at 33-42, which identified "six broad categories [of civil liberties risks]: (1) chilling effect and other surveillance risks; (2) data aggregation; (3) data inaccuracy; (4) data misuse; (5) false positives; and (6) risks associated with data processing." Id. at 35.

n53 Camara v. Municipal Court, 387 U.S. 523, 528 (1967) (safety inspectors could not enter home without warrant, but need not demonstrate probable cause); Wolf v. Colorado, 338 U.S. 25, 27 (1949) ("The security of one's privacy against arbitrary intrusions by the police—which is at the core of the Fourth Amendment—is basic to a free society.").

n54 "Absent some grave emergency, the Fourth Amendment has interposed a magistrate between the citizen and the police. This was done . . . so that an objective mind might weigh the need to invade [the citizen's] privacy in order to enforce the law." McDonald v. United States, 335 U.S. 451, 455 (1948); see Steagald v. United States, 451 U.S. 204, 212 (1981) (warrants needed because police "may lack sufficient objectivity"). Even when the police conduct a search properly, accountability matters. Katz v. United States, 389 U.S. 347, 356 (1967) ("the inescapable fact is that this restraint was imposed by the agents themselves, not by a judicial officer.").

n55 Maryland v. Garrison, 480 U.S. 79, 84 (1987) ("By limiting the authorization to search to the specific areas and things for which there is probable cause to search, the requirement ensures that the search will be carefully tailored to its justifications," so that "the scope of a lawful search is defined by the object of the search and the places in which there is probable cause to believe that it may be found.") (internal quotation marks and citation omitted); Lo-Ji Sales, Inc. v. New York, 442 U.S. 319, 325 (1979) ("Nor does the Fourth Amendment countenance open-ended warrants, to be completed while a search is being conducted and items seized or after the seizure has been carried out.").
Importantly, Groh also makes clear that the purpose of the particularity requirement is not limited to the prevention of general searches. Id. ("A particular warrant also assures the individual whose property is searched or seized of the lawful authority of the executing officer, his need to search, and the limits of his power to search," which "greatly reduces the perception of unlawful or intrusive police conduct") (internal quotation marks and citations omitted).

Chandler v. Miller, 520 U.S. 305, 308 (1997) (Fourth Amendment "generally bars officials from undertaking a search or seizure absent individualized suspicion" except in "certain limited circumstances"); but see id. at 323 ("where the risk to public safety is substantial and real, blanket suspicionless searches calibrated to the risk may rank as 'reasonable'-for example, searches now routine at airports and at entrances to courts and other official buildings"); see also id. at 313-314 ("special needs, beyond the normal need for law enforcement," may justify departures) (citation omitted). Whether counterterrorism investigation fits this "special needs" category is beyond the scope of this essay.

Probable cause exists if the facts and circumstances known to the officer warrant a prudent man in believing that the offense has been committed.” Henry v. United States, 361 U.S. 98, 102 (1959).

The fact that the application adequately described the 'things to be seized' does not save the warrant from its facial invalidity. The Fourth Amendment by its terms requires particularity in the warrant, not in the supporting documents.” (emphasis in original).

The general warrant, in which the name of the person to be arrested was left blank, and the writs of assistance . . . both perpetuated the oppressive practice of allowing the police to arrest and search on suspicion. Police control took the place of judicial control, since no showing of 'probable cause' before a magistrate was required.” Henry, 361 U.S. at 100 (footnote omitted); see Katz v. United States, 389 U.S. 347, 358- 359 (1967) ("bypassing a neutral predetermination of the scope of a search leaves individuals secure from Fourth Amendment violations only in the discretion of the police") (internal quotation marks and citation omitted).


Clancy, supra note 15, at 497-505.

Chandler, 520 U.S. at 313 ("a search ordinarily must be based on individualized suspicion of wrongdoing.").
n65 Berger, 388 U.S. at 59.


n67 388 U.S. 41 (1967).

n68 Id. at 54-55.

n69 Id. at 58.

n70 Id. at 56.


n72 Id. at 88 (internal quotation marks omitted).

n73 Id. at 88-91 (citation omitted). The warrant did not authorize the search of patrons, and the warrant application did not mention tavern patrons at all. Id. at 90 n.2. Nor did Ybarra's behavior during the search give the police any reason to believe he was engaged in illegal activity. Id. at 90-91.

n74 Id. at 91.

n75 TAPAC Report, supra note 11, at 45.

n77 Berger, 388 U.S. at 57.

n78 Id.

n79 Id.


n81 Taipale, supra note 13, at 64.

n82 Tether testimony, supra note 14.

n83 William J. Stuntz, Local Policing After the Terror, 111 Yale L.J. 2137, 2167 (2002) (arguing that "sweep" searches of large groups are "more likely to occasion public complaint, and hence less likely to involve the kind of behavior that prompts complaint . . . .").

n84 "The progress of science in furnishing the Government with means of espionage is not likely to stop with wire-tapping. Ways may some day be developed by which the Government, without removing papers from secret drawers, can . . . expose to a jury the most intimate occurrences of the home." Olmstead v. United States, 277 U.S. 438, 474 (1928) (Brandeis, J., dissenting).

n85 The Court has noted the danger of relying only on after-the-fact review. See Katz v. United States, 389 U.S. 347, 358 (1967) ("after-the-event justification . . . [is] too likely to be subtly influenced by the familiar shortcomings of hindsight judgment.") (internal quotation marks and citation omitted); id. at 356-357 (the Court "has never sustained a search upon the sole ground that officers reasonably expected to find evidence of a particular crime and voluntarily confined their activities to the least intrusive means consistent with that end.").

n86 United States v. U.S. District Court, 407 U.S. 297, 318 (1972) ("post-surveillance review would never reach the surveillances which failed to result in prosecutions. Prior review by a neutral and detached magistrate is the time-tested means of effectuating Fourth Amendment rights.") (citation omitted).
n87 Id. at 317 (footnote omitted).

n88 291 F. Supp. 2d 1132 (C.D. Cal. 2003). Garcetti was a federal § 1983 action brought by three groups of plaintiffs challenging the “hand off” procedure described below.

n89 Id. at 1136-1137.

n90 Id. at 1138.

n91 Id. at 1139 n.14.

n92 Id. at 1138-1139.

n93 Whitaker, 291 F. Supp. 2d at 1146; id. at 1148 (”The Court cannot imagine a more effortless nullification of the probable cause requirement than a concealment of the existence of the affidavit, the wiretapping order, and the resulting wiretap.”).

n94 Neural networks are an approach to computing based loosely on the architecture of animal brains, e.g., networks of processing elements or nodes analogous to neurons. These nodes are interconnected in a network that can learn to identify patterns as it is exposed to data.

n95 Two Crows, supra note 22, at 11.

n96 Id.

Christopher Slobogin argues, somewhat similarly, that public video surveillance should be subject to Fourth Amendment regulation modeled on the Court's approach to automobile checkpoints. Christopher Slobogin, Public Privacy: Camera Surveillance of Public Places and the Right to Anonymity, 72 Miss. L.J. 213 (2002).

Zarsky, supra note 20.


Jensen, supra note 24, at slide 21.

Cf., Skinner v. Railway Labor Executives' Ass'n, 489 U.S. 602, 616 (1989) (when obtaining physical evidence from a person, initial detention may be a seizure, while obtaining and examining the evidence may be a search).


The key question is whether the sender retains control over the item and its contents. See, e.g., United States v. Most, 876 F.2d 191, 197-98 (D.C. Cir. 1989) (finding reasonable expectation of privacy in contents of plastic bag left with grocery store clerk); United States v. Barry, 853 F.2d 1479, 1481-83 (8th Cir. 1988) (finding reasonable expectation of privacy in locked suitcase stored at airport baggage counter); United States v. Presler, 610 F.2d 1206, 1213-14 (4th Cir. 1979) (finding reasonable expectation of privacy in locked briefcases stored with defendant's friend for safekeeping).


United States v. Walser, 275 F.3d 981, 986 (10th Cir. 2001); see United States v. Carey, 172 F.3d 1268, 1273-75 (10th Cir. 1999) (agent exceeded scope of warrant to search for evidence of drug sales when he "abandoned that search" and instead searched for evidence of child pornography for five hours); but see United States v. Slanina, 283 F.3d 670, 680 (5th Cir. 2002) (justified warrantless search of a portion of computer and storage device destroyed defendant's reasonable expectation of privacy in remaining contents); United States v. Runyan, 275 F.3d 449, 464-65 (5th Cir. 2001).

n108 Id. at 651-652.

n109 Id. at 659 (footnote omitted).


n111 Id. at 338-339.

n112 Id. at 337; id. at 339 ("the agent's physical manipulation of petitioner's bag violated the Fourth Amendment.").

n113 In Bond, the Court noted that "petitioner sought to preserve privacy by using an opaque bag and placing that bag directly above his seat." Id. at 338.

n114 In Walter, the employee who held the films up to the light could not make out the images because "the film itself is only four millimeters wide," and "[s]ince the scenes depicted . . . are necessarily even more minute, it is easy to understand why such films cannot be examined successfully with the naked eye." Walter, 447 U.S. at 652 n.2.

n115 In Whalen v. Roe, 429 U.S. 589 (1977), the Court said that it was "not unaware of the threat to privacy implicit in the accumulation of vast amounts of personal information in computerized data banks or other massive government files," expressed concern for potential abuse of that information, and suggested that there may be a constitutional duty to avoid unwarranted disclosure. Id. at 605 (footnote omitted).

n117 Id. at 763-764. The Court further noted that "[m]eaningful discussion of privacy . . . requires the recognition that ordinarily we deal not with an interest in total nondisclosure but with an interest in selective disclosure." Id. at 763 n.14 (quoting Kenneth Karst, "The Files": Legal Controls Over the Accuracy and Accessibility of Stored Personal Data, 31 Law & Contemp. Probs. 342, 343-344 (1966)). The Court was, however, careful to distinguish "the question whether an individual's interest in privacy is protected by the Constitution." Id. at 762 n.13.

n118 E.g., State v. Von Bulow, 475 A.2d 995, 1018 (R.I. 1984) (in murder case where death was allegedly caused by administration of drugs, warrantless toxicological testing that exceeded scope of private search and "positively identified the unknown composition of the pills" "by employing chemical or mechanical means to reveal the[s] hidden nature" violated Fourth Amendment); cf. State v. Joyce, 639 A.2d 1007, 1015 (Conn. 1994) (warrantless chemical analysis "capable of determining a multitude of private facts about an individual" that detected gasoline in arson suspect's clothing was search under state constitution) (footnotes omitted).


n120 Id. at 617.


n122 Id. at 122.

n123 Id. at 123 (footnote omitted); see United States v. Place, 462 U.S. 696, 707 (1983) (dog sniff for marijuana offends no legitimate privacy interest because detects only presence or absence of contraband).


n125 In a very convoluted way, the Court struggled with the issue of cognition enhancement in Kyllo v. United States, 533 U.S. 27 (2001). The question in Kyllo was whether police use of a thermal imaging device to measure the heat emanating from a house was a search. The dissent argued that the thermal imager merely observed "the exterior of the home," as opposed to "through-the-wall surveillance' that gives the observer or listener direct access to information in a private area." Id. at 41 (Stevens, J., dissenting). Thus, the case involved "nothing more than drawing inferences from off-the-wall surveillance . . . ." Id. at 46 (Stevens, J., dissenting). Indeed, the dissent claimed that the majority "effectively treats the mental process of analyzing data obtained from external sources as the equivalent of a physical intrusion into the home." Id. at 49 (Stevens, J., dissenting). The majority, on the other hand, argued that neither type of surveillance was truly "direct," and that both involved "inferences." Id. at 36 ("Surely the dissent does not believe that the through-the-wall radar or ultrasound technology..."
produces an 8-by-10 Kodak glossy that needs no analysis (i.e., the making of inferences).”). The majority agreed that "an inference is not a search," but that "hi-tech measurement of emanations from a house is a search." Id. at 37 n.4.


n128 See, e.g., id. at 764 (distinguishing "scattered disclosure of bits of information contained in a rap sheet and revelation of the rap sheet as a whole."); Nader v. General Motors Corp., 307 N.Y.S.2d 647, 657 (N.Y. 1970) ("arguably a right to privacy may . . . be invaded through extensive or exhaustive monitoring and cataloguing of acts normally disconnected and anonymous.") (Breitel, J., concurring); Jerry Kang, Information Privacy in Cyberspace Transactions, 50 Stan. L. Rev. 1193, 1260-61 (1998) (distinguishing "casual observation" from "surveillance").


n130 Id. § 2510(4) (defining "interception" as the use of a device to acquire the contents of a communication).

n131 Walker v. Darby, 911 F.2d 1573, 1578-79 (11th Cir. 1990) ("it is highly unlikely that he would have expected his conversations to be electronically intercepted and monitored in an office in another part of the building."); see Erving Goffman, Relations in Public 286 (1971) ("A working assumption in everyday life is that one's surroundings will be 'dead,' that is, contain no recording and transmission devices.").

n132 LaFave, supra note 16, at 309 (footnote omitted) (emphasis omitted).