# II. THE IRS

The IRS is the branch of the United States Department of Treasury that is responsible for administering the Internal Revenue Code and enforcing tax law. Income taxes were introduced to the United States in 1913 when the Sixteenth Amendment was enacted. While the Treasury Department collects the taxes, the IRS is responsible for examining the tax returns for accuracy and bringing criminal action against those who file incorrect returns. Each tax return is checked internally for mathematical accuracy and consistency, regardless of whether it is submitted via mail or electronically. The IRS also compares the submitted returns to third-party materials that are required to be filed with the IRS, such as W-2s and 1099s. Today the IRS is taking advantage of the large amount of data that can be purchased from data brokers as well as amassing its own data sets.

## A. IRS Data Collection

Prior to discussing the potential issues with the IRS’s use of data analytics, it is important to understand what data it is collecting and from where it is collecting that data. While the IRS may request information from taxpayers to support the information provided on their tax returns, individuals are having to consider the constitutionality of the IRS collecting and maintaining information on taxpayers from sources other than the taxpayer and *prior*to an audit. Even though a taxpayer is required to maintain the proof necessary to support any line item on a tax return, the taxpayer need not provide support along with her return, nor would she need to support an allowed deduction, such as the payment of mortgage interest, if she instead chose to take the standard deduction or simply not take the deduction at all. While the burden is on the taxpayer to support their return, the IRS does not have unlimited power to obtain any data it desires regarding a taxpayer.

It is well known that the IRS is able to obtain information from third parties to verify line items on tax returns provided by taxpayers. An example would be a W-2 from an employer. However, the right to third party information is not unlimited. Only recently have privacy scholars begun to examine these issues when it comes to electronic and phone communications. Most of the rules permitting the IRS to obtain records from third parties were written prior to the existence of social media, and certainly prior to the current state of technology. “Modern technologies are creating ‘minutely detailed records’ of our existence, increasingly facilitating the ‘persistent, continuous and indiscriminate monitoring of our daily lives.’” The existence of data brokers and the ability to purchase information about pretty much anyone over the Internet has created a situation where users are losing control over who sees their once private information. This is especially unsettling when that viewer is the IRS.

### 1. Phone Records

According to the American Civil Liberties Union (ACLU), the IRS is one of the agencies that purchased cell phone tracking technology in 2009 and 2012. This phone tracking technology, known as Stingray, masks as a cell tower to trap metadata and content from cell phones that connect to them. This technology means the IRS has the ability to record phone conversations, text messages, and track the location of individuals using their cell phones without anyone being aware of this tracking. Legal scholars believe that the IRS will increasingly rely on surveillance technology to reduce noncompliance. A case is currently being heard in Maryland regarding the constitutionality of the government’s use of the Stingray cell tracking device. Although in 2015 the Department of Justice (DOJ) issued a guidance statement for the department’s law enforcement constituents, these guidelines do not apply to the IRS.

### 2. Emails

Pursuant to a Freedom of Information Act (FOIA) request in 2013, the ACLU discovered that the IRS had been reading taxpayers’ private emails without a warrant. The 2011 IRS auditor’s training manual indicated that investigators could obtain everything in an account using an Electronic Communications Privacy Act (ECPA) court order except for unopened email or voicemail stored with a provider for 180 days or less. This policy is in direct contravention of the 2010 ruling in *United States v. Warshak*, which reaffirmed that citizens have a reasonable expectation of privacy in their emails and that the government needs a warrant to obtain them. It should be noted that an ECPA court order can be issued fairly easily and does not require “probable cause” that a criminal statute has been violated. In response to a Senate Finance Committee hearing, the IRS agreed to stop reading taxpayers’ emails without a warrant but was notably silent about its social media activities.

### 3. Social Media

According to a spokesperson for the UC-Berkeley Samuelson Clinic, the IRS confirmed in response to a FOIA request that it is collecting information from social media sites. An IRS training document mentions Facebook, MySpace, and YouTube as possible sources for taxpayer information. According to CNET, the IRS uses “online activity trackers to look though mass amounts of public Internet data for potentially incriminating information.” The IRS has also used evidence from Google Maps in a Tax Court case to revoke the 501(c)(4) tax exempt status of a homeowners’ association. There is, of course, a difference between locating publicly available information online about a taxpayer who is being audited and data mining for potential tax violators prior to the time the taxpayer has been selected for an audit. The IRS is reported to have used automated computer programs (sometimes known as spiders) to sort through social media sites.

### 4. Data Mining

Data mining involves the analysis of large data sets, which have been collected for a purpose other than that for which they are being analyzed, in order to search the data sets for previously unknown relationships in the data. Data mining can be descriptive or predictive: descriptive data mining summarizes properties of the data set, while predictive data mining performs analysis on a data set to build a model that makes predictions about data that is not available. The IRS engages in data mining in order to develop analytics and algorithms to identify tax compliance issues. According to an IRS report: “It is not possible to have compliance experts review every possible set of related tax returns. . . . Active learning can be used to refine targeting models. Common connections between possibly abusive transactions can be used to identify potential promoters of these transactions.” Sources have disclosed that the IRS is using data mining to create more detailed profiles of taxpayers. “If Nike is analyzing my information, the worst consequence is that they market stuff to me that I don’t want and it’s annoying,” stated Behnam Dayanim, co-chair of the privacy and data practice at Paul Hastings, “[i]f the government does it, the worst consequence is there could be legal ramifications, whether it’s fines, penalties, or imprisonment.” Concerns about agency use of data mining were also discussed in the Senate hearings regarding the Federal Agency Data Mining Reporting Act of 2007. The hearing report indicated that there were 199 different government data mining programs, including the IRS, and that there was very little control over these activities.

## B. History of Improper Audits

One of the concerns with the IRS’s unprecedented access to private information is the IRS’s history of misusing the audit function. An audit is intended to ensure tax compliance; the IRS audits returns to check for mathematical errors, document mismatching, and noncompliance. However, since the creation of the IRS, government officials, particularly presidents, have been using the IRS for their own political agendas.

President Franklin D. Roosevelt set the stage for presidents using the IRS as a weapon to investigate political rivals and business opponents. Roosevelt’s victims included Senator Huey Long, United Mine Workers leader John Lewis, Representative Hamilton Fish, Chicago Tribune publisher Robert “Colonel” McCormick, Philadelphia Inquirer publisher Moses Annenberg, William Randolph Hearst,

Father Charles Coughlin, and, the former Treasury Secretary Andrew Mellon. The latter is especially ironic because Andrew Mellon utilized the IRS to audit his rivals as Treasury Secretary under President Calvin Coolidge.

Between 1956 and 1971, the FBI ran a counterintelligence program called COINTELPRO. The brainchild of then-FBI Director J. Edgar Hoover, its purpose was initially to disrupt, discredit, and destroy Communist Party activities in the United States. It later expanded to include other groups such as the Socialist Workers Party and the Black Panther Party. Under COINTELPRO, the FBI was able to harass these individuals and organizations by having the IRS target them for tax audits. Martin Luther King Jr. was a victim of this harassment, as was the National Association for the Advancement of Colored People and the National Council of Churches.

When Robert Kennedy, chief counsel for the Senate Select Committee on Improper Activities in Labor and Management, investigated Teamsters leader Jimmy Hoffa in the late 1950s for illegal activities, he failed to obtain a conviction. When he was the Attorney General under President John F. Kennedy, Robert again sought out Jimmy Hoffa. One of the tactics used was requesting the IRS to repeatedly examine his returns and those of his associates for tax evasion. This tactic was also employed for other alleged racketeers whom Kennedy had his eye on. This targeting of those believed to be involved in criminal activity raised questions from legal experts who decried that tax laws are for revenue collection, not prosecuting criminals, and insisted that audits should be random. Under the Kennedy administration, IRS investigations extended to groups with extreme conservative views such as the John Birch Society. The IRS went so far as to establish the “Ideological Organizations Audit Project” to target these groups.

The President who really excelled at wielding the IRS audit weapon against political enemies was Nixon. Besides targeting left-wing groups, Nixon sought out antiwar groups, churches and nonprofits supporting antiwar groups, civil rights groups, reporters, and prominent Democrats. The White House tapes provide direct evidence of Nixon using the IRS to collect data on potential Democratic presidential candidates, including Senators Hubert

Humphrey, Edward (Ted) Kennedy, and Edmund (Ed) Muskie. Nixon had the IRS establish the Special Service Staff unit to utilize tax records to create dossiers on more than 11,000 individuals and groups, including supporters of Democrat Presidential nominee George McGovern for 1972. In the House Judiciary Committee’s 1974 Articles of Impeachment, one of the articles charged Nixon with trying to obtain confidential information contained in income tax returns for purposes not authorized by law, in violation of the taxpayer’s constitutional rights, and causing the selection of audits in a discriminatory manner. There have also been IRS abuses by presidents subsequent to Nixon. Recently, the IRS singled out conservative organizations with “tea party” affiliations that were seeking tax-exempt nonprofit status and subjecting them to extra scrutiny. While there may or may not have been a political motivation for the increased scrutiny, it does appear that at the very least gross mismanagement was involved. Despite the Department of Justice’s finding that no criminal conduct occurred, the House Ways and Means Committee has indicated that it will continue to investigate the targeting.

## C. Audit Selection History

Every year the IRS must shift through copious numbers of taxpayer returns and their related data. In order to ensure tax compliance, the IRS may audit a tax return to check for mathematical errors, document mismatching, and noncompliance. The audit may be performed through the mail, at the taxpayer’s home, or at an IRS office. Historically, tax returns were selected randomly (based on a statistical formula), due to a mismatch with third party data, or when a return was linked to other taxpayers who were being audited themselves. The majority of audits resulted from mismatches with third party data. The Information Returns Processing (IRP) System was responsible for the data received from employers and other third parties reporting taxpayer income, pensions, interest, and dividends paid during the tax year. The IRP would match income reported on information returns against income reported by taxpayers on their individual income tax returns based on Social Security numbers. When mathematical errors, inconsistencies, or a mismatch in the IRP system was identified, the taxpayer was contacted via mail and a bill or check was sent to the taxpayer.

The rate of audits increased during the 1950s. In the early 1960s over 5.5 percent of tax returns were chosen for audits. The rate then began declining due to technological advancements in identifying potential tax returns to audit. The IRS first used computers for selecting tax returns in 1962 and created the Taxpayer Compliance Measurement Program (TCMP) two years later. The TCMP randomly selected about 50,000 returns approximately every three years to perform detailed audits requiring substantiation of each line on the tax return. This program initially reviewed delinquent returns to create a statistical summary, which evolved into an automated program known as the discriminant function analysis (DIF). The DIF gives each tax return a score based on the probability of noncompliance. IRS personnel then manually screen the tax returns to ensure appropriate selection. This process enhanced audit efficiency by allowing the IRS to manually review the machine scored returns and chose the tax returns with the highest likelihood of noncompliance while avoiding auditing compliant returns. The first tax audits based on the DIF occurred in 1969, and refinements to the DIF were made during the 1970s and again in the 1980s with the addition of computerized third-party document matching and mathematical accuracy. Analysis of the most common errors by taxpayers led to policy changes. For example, in 1986 the TCMP identified a significant misreporting of dependency exemptions and wrongful claims of the earned income credit by individuals claiming children that did not qualify. The policy was then changed to require identification numbers for dependents. As a result, the number of dependents claimed in 1987 was 7 million fewer than claimed in 1986 when identification numbers were not required. A similar decrease was found for those claiming an earned income credit. In the 1980s, when third party reporting became required of income items such as wages, interest, and dividends, the accuracy of these amounts on tax returns substantially increased.

Prior to the time the TCMP audit was used, only half of the audited returns found any errors. During the time of the TCMP audits, the percentage of returns chosen for audits containing no errors (no-change) decreased from over 40 percent in 1968 to about 11 percent in the early 1990s. However, TCMP audits were onerous because they required the taxpayer to support each line of their tax return with documentation. In 1988, the TCMP was eventually phased out due to cuts in the IRS budget and criticisms by taxpayers, Congress, and the media. The 1996 General Accounting Office (GAO) report on the IRS, suggested that the IRS find alternative methods (to the TCMP) for updating the DIF and develop a long-term strategy for obtaining compliance data with fewer resources. Supporting the GAO’s predictions of the detrimental effects of not updating the TCMP, the 1994 no-change rate for individual returns identified by the DIF was over 19 percent, and more than 24 percent in 1998. By not updating the data upon which the DIF was based, a greater percentage of compliant taxpayers had to suffer through the expense and stress of an audit. Based on concern that the effectiveness of the DIF was deteriorating and reducing taxpayer confidence in the fairness of the tax system, in 2002 the IRS initiated the National Research Program (NRP) to replace the TCMP. The idea was to increase the quality of the data and better predict which tax returns would result in a deficiency without the burdensome TCMP audits.

The NRP gathers data learned from random audits to measure voluntary compliance with tax laws and improve DIF audit selection methods. This data is used for analytical purposes such as identifying tax issues, reporting characteristics, and taxpayer segments that may lead to noncompliant behavior. The rationale for performing this type of analytics is that efficiency in the audit process reduces unnecessary audits for compliant taxpayers. Today, the IRS uses big data analytics to target their audits on tax returns more likely to result in tax deficiencies.

Becoming more efficient has become increasingly important as the budget for the IRS continues to decrease. Since 2010, the budget has been cut by 17 percent and staff has decreased by 14 percent. This has caused a decrease in compliance monitoring. In fact, the number of audits dropped to an eleven-year low in 2015, in which the IRS collected $3.3 trillion in revenue and processed about 243 million tax returns. This corresponded to 35 cents spent for each hundred dollars it collected. Although the IRS is one of the world’s most efficient tax administrations, IRS Commissioner John Koskinen stated that there comes a point when it is not possible to keep doing more with less without jeopardizing the mission of the IRS. He also projected that with a larger budget, the IRS could increase the amount of taxes collected. In his written statement, Koskinen added, “I don’t know any organization in my 20 years of experience in the private sector that has said, ‘I think I’ll take my revenue operation and starve it for funds to see how it does.’” It is estimated that for every dollar decrease in the IRS budget, there are five dollars owed that are not collected.

Most of the changes the IRS has made to address their budget shortfall rely on the increased use of technology. For the 2015 tax return filing season, around 90 percent of the returns were filed electronically, thus reducing the need for data entry employees. In 2005 electronic filings comprised only 50 percent of the total. The IRS uses the Automated Under-Reporter Program to match third-party information reports with tax returns and contact taxpayers via letters to resolve discrepancies. In 2015, the IRS received 2.6 billion third-party information reports, of which over 87 percent were filed electronically. The ability to easily verify tax return information allowed the IRS to resolve more than 3.7 million tax return discrepancies in 2015, resulting in more than $6.3 billion. All of this was accomplished with the equivalent of only 1,739 fulltime employees—approximately a $3.6 million increase in post-audit collections per employee.

Data analytics is being touted as the solution to the IRS’s budget problems. Part of the IRS’s data analytics program examines source data to identify noncompliant tax returns going beyond information provided by the taxpayer and the third-party sources required to submit information (such as employers providing W-2s). The IRS asserts that with analytics it can improve efficiencies and effectiveness of its investigations and avoid wasting taxpayers’ time or creating unnecessary burdens on them. Koskinen opined that without analytics, the future of the IRS would not be possible. According to Dean Silverman, at that time the IRS’s senior adviser to the commissioner for the Office of Compliance Analytics, the IRS is expanding its source data resources to include credit and debit card processors, PayPal, social media, and other Internet data.