Kimberly A. Houser

Assistant Professor

Oklahoma State University

NBC NEWS **SPECIAL REPORT**

LIVE

FACEBOOK CEO MARK ZUCKERBERG TESTIFIES BEFORE SENATE

NBC NEWS

MR. HATCH

THE IRS AND THE DATA MESS

Kimberly A. Houser, Assistant Professor, Oklahoma State University

Office of Compliance Analytics

Due to a decrease in personnel
and an increase in filed tax returns,
the IRS has turned to big data analytics
to solve their budget woes.
The Office of Compliance Analytics
was created in 2011 as a new division
of the IRS.
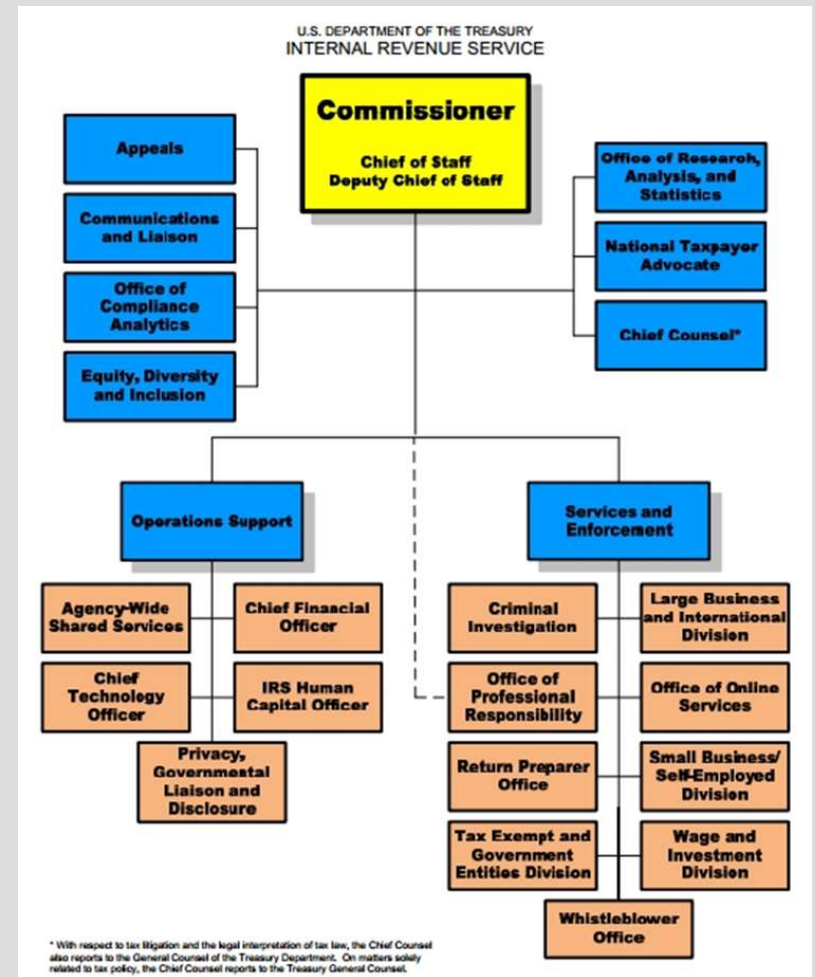
IRS Data Mining

It has been reported that the IRS is now mining commercial and public data pools including social media sites (e.g., Facebook, Instagram, and Twitter) and pattern recognition algorithms are then run on the data to identify potential non-compliant taxpayers.

SPOKESMAN.COM   SPECIAL REPORT

IS THE TAXMAN CYBERSTALKING YOU? YES.

We are at a time when we can guide the future of privacy, data security, and AI

The focus on private industry's use of data neglects the real risk of uncontrolled government surveillance

However, such laws must apply equally to government and private industry

Government has been mining the internet for information and using predictive analytics to make decisions impacting human rights

There is an enormous difference between using an algorithm to predict which book you like and how long you should stay in jail

TEST

ROBOTRON EC8840

Data Collection - The Privacy Act of 1974
The Privacy Act limits the federal government's use of private data about US citizens

Data Minimization - CMPPA of 1988
The Computer Matching and Privacy Protection Act requires that each agency "maintain in its records only such information about an individual as is relevant and necessary to accomplish a purpose of the agency

Data Quality – PRA of 1995
The Paperwork Reduction Act requires federal agencies to take steps to ensure the quality of their data

Data Mining/Analytics Examples

Although the IRS website in no way reveals this to taxpayers, Dean Silverman, former Senior Advisor to the Commissioner in the Office of Compliance Analytics for the Internal Revenue Service, indicated that the IRS uses big data for the following:

• Charting and analyzing social media such as Facebook
• Targeting audits by matching tax filings to social media or electronic payments
• Tracking individual Internet addresses and emailing patterns
• Sorting data in 32,000 categories of metadata and 1 million unique "Attributes"
• Machine learning across "neural" networks
• Statistical and agent-based modeling
• Relationship analysis based on Social Security numbers and other personal identifiers

POTENTIAL HARMS

1. Discrimination
2. Misuse of audit function
3. Data breach

Potential Discrimination

Data mined from the internet or purchased from data brokers will contain historical and societal bias

Predictive analytics has the potential to use a proxy for race and religion
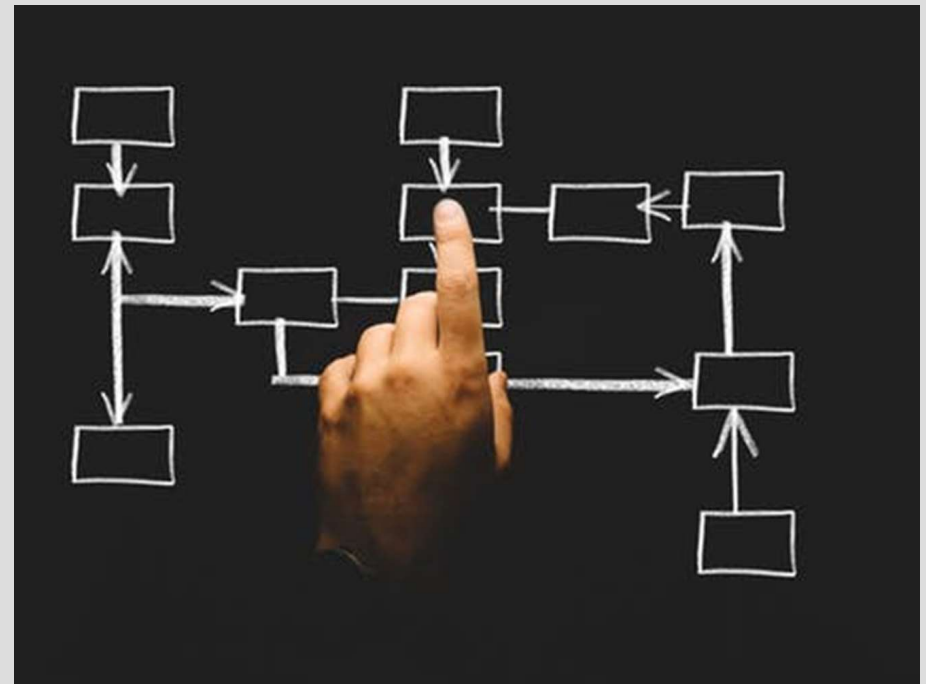
The government has a long history of misusing the IRS audit function for political purposes

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, but goes back all way to Roosevelt.

Targeting by Government

During the 2012 presidential election,
the IRS started flagging conservative
political groups for additional reviews to
see if they were violating their tax-exempt status.

DATA BREACH

1. The treasury inspector general reported that 21% of FOIA requests answered by the IRS wrongly disclosed "sensitive taxpayer information."

2. The IRS reported that more than 720,000 social security numbers had been stolen from the "get transcript" function on its website.

3. Hackers were also able to obtain 100,000 e-pin numbers to claim fraudulent refunds.

A 2015 audit GAO of security procedures listed 43 deficiencies and revealed that the IRS had ignored previous audit recommendations and was not keeping taxpayer data secure.

GAO

Highlights

Highlights of GAO-15-337, a report to the Commissioner of Internal Revenue

INFORMATION SECURITY

IRS Needs to Continue Improving Controls over Financial and Taxpayer Data

Why GAO Did This Study

The IRS has a demanding responsibility in collecting taxes, processing tax returns, and enforcing the nation's tax laws. It relies extensively on computerized systems to support its financial and mission-related operations and on information security controls to protect the financial and sensitive taxpayer data that resides on those systems.

As part of its audit of IRS's fiscal year 2014 and 2013 financial statements, GAO assessed whether controls over key financial and tax-processing systems were effective in ensuring the confidentiality, integrity, and availability of financial and sensitive taxpayer information. To do this, GAO examined IRS information security policies, plans and procedures; interviewed key agency officials; and tested controls over key financial applications at four sites.

What GAO Recommends

GAO is recommending that IRS take 5 additional actions to more effectively implement elements of its information security program. In a separate report with limited distribution, GAO is recommending 14 actions that IRS can take to address newly identified control weaknesses. In commenting on a draft of this report, IRS agreed to develop corrective action plans where appropriate to address these recommendations.

What GAO Found

The Internal Revenue Service (IRS) made progress in implementing information security controls; however, weaknesses limit their effectiveness in protecting the confidentiality, integrity and availability of financial and sensitive taxpayer data. During fiscal year 2014, IRS continued to devote attention to securing its information systems that process sensitive taxpayer and financial information. Key among its actions were improving the security over the software that manages changes to its mainframe environment and upgrading secure communications enterprise-wide for sensitive data. However, significant control deficiencies existed. For example, IRS did not install appropriate security updates on all of its databases and servers, and did not sufficiently monitor control activities that support its financial reporting. In addition, IRS did not effectively maintain the secure configuration of a key application, or appropriately segregate duties by allowing a developer unnecessary access to the application.

An underlying reason for these weaknesses is that IRS has not effectively implemented elements of its information security program. The agency had a comprehensive framework for its program, such as assessing risk for its systems, developing security plans, and providing employees with security awareness and specialized training. However, aspects of its program were not yet effectively implemented. For example, IRS's testing methodology did not always determine whether required controls were operating effectively; consequently, GAO continued to identify control weaknesses that had not been detected by IRS. Also, IRS had not updated key mainframe policies and procedures to address issues such as comprehensively auditing and monitoring of access, thereby increasing the risk of unauthorized access to tax processing systems not being detected. In addition, IRS did not reassess controls for a key system after significant changes had been made in the operating environment. Further, IRS had not ensured that many of its corrective actions to address previously identified deficiencies were effective. For example, of 69 previously reported weaknesses that remained unresolved at the end of GAO's last audit, IRS indicated it had implemented corrective actions for 24 of them; however, GAO determined that 10 of the 24 weaknesses had not been fully resolved.

Until IRS takes additional steps to (1) address unresolved and newly identified control deficiencies and (2) effectively implements elements of its information security program, including, among other things, updating policies, test and evaluation procedures, and remedial action procedures, its financial and taxpayer data will remain unnecessarily vulnerable to inappropriate and undetected use,

The IRS's use of big data raises issues of data quality, privacy, politics, and race

- The method of data collection (purchasing data sets and mining social media) is suspect and does not comply with FIPs.

- The maintaining of such information beyond reviewing an individual tax return violates the law.

- The more information the IRS maintains about individuals, the easier it becomes to target audits to political rivals.

- The use of algorithms can result in discrimination by secretly identifying race as a tax cheat risk.

- The collection of such highly detailed information by the IRS makes it a very attractive target for hackers.
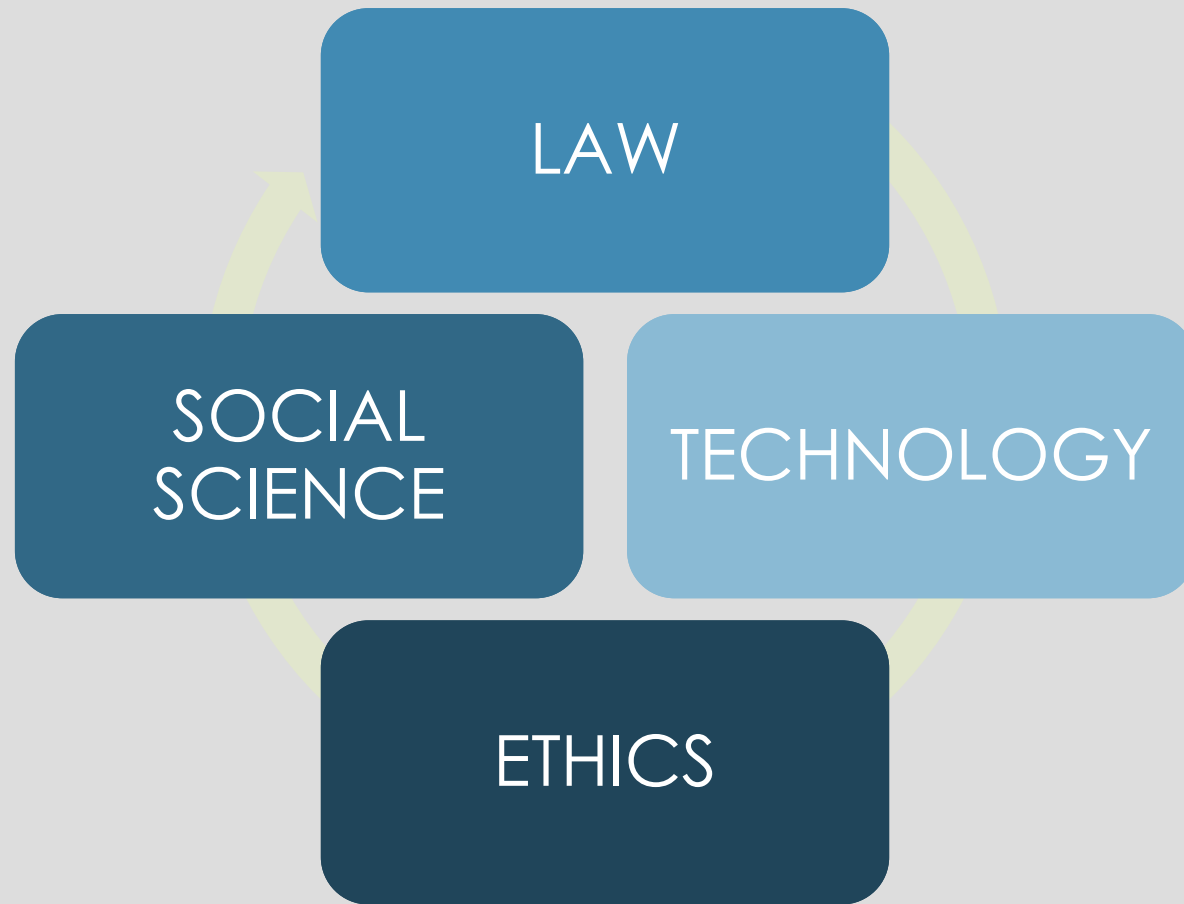
New Research

1. How AI in employment decisions can mitigate bias and noise resulting from human decision-making and fix the diversity problem in the tech industry (forthcoming 2019)

2. Why China will win the AI race – a comparison of U.S., EU and Chinese law

3. Reducing the pay gap in the U.S. through the use of artificial intelligence

4. How micro-targeting discriminates in the private and public sector

5. The top 10 reasons to be optimistic about AI

Any Questions?

Houser, K.A. and Voss, G. (Dec. 20, 2018). The European Commission on the Privacy Shield: All Bark and No Bite?. *Illinois Journal of Law, Technology & Policy – Timely Tech*.

Houser, K. A. and Voss, G. (2018). GDPR: The end of Google and Facebook or a New Paradigm in Data Privacy? *Richmond Journal of Law & Technology 25-1*.

Houser, K.A. and Voss, G. (2018, August 29). Can Facebook and Google survive the GDPR? *Oxford Business Law Blog* reprinted in *Sage Ocean London Guest Blog* (2018, Sept. 26).

Houser, K. A. & Sanders, D. (2018, Feb.). The use of big data analytics by the IRS: what tax practitioners need to know. *Journal of Taxation, 128-2*.

Houser, K.A. (2018, Jan. 20). IRS violating privacy law – must do better. *The Hill*.

Houser, K.A. (2017, Sept. 28). Clear Social Media Policies are a Must. *Spokane Journal of Business*.

Houser, K. A. & Sanders, D. (2017). The use of big data by the IRS: efficient solution or the end of privacy as we know it? *Vanderbilt Journal of Entertainment & Technology Law, 19-4*.

Chen, L., Gramlich, J., and Houser, K. A. (2017). The effects of board gender diversity on a firm's financial and reputation risk strategy. *Accounting and Finance, 57-3*.