

SPONSORED BY:

MOBILE



Transform Your Business With Data

[ABOUT US](#) [REGISTER](#) [LOGIN](#)

[HOME](#) [BLOGGERS](#) [MESSAGES](#) [POLLS](#) [WEBINARS](#) [RESOURCES](#) [VIDEOS](#) | [STRATEGY](#) [ANALYTICS](#) [TECHNOLOGY](#) [FB](#) [TW](#) [LI](#) [G+](#) [RSS](#)

[FINANCIAL SERVICES](#) [HEALTHCARE](#) [GOVERNMENT](#) [EDUCATION](#) [RETAIL](#)

BIG DATA ANALYTICS

The IRS: Big Data or Big Brother?



Robert Plant, Associate Professor, School of Business Administration, University of Miami
 4/17/2013
 Comment
 19 comments

[Login](#)
 50% 50%

You may have spent the last few days thinking hard about your taxes, but the IRS has been doing so for years -- positioning itself as a leader in pragmatic big data use.

Each year, April 15 is a memorable date for those of us in the United States -- this is the deadline to file our taxes. The IRS is, without question, the dominant government agency in the US. No taxes, no government -- simple. And thus generations of politicians since independence have ensured that the IRS has all the powers it needs to raise federal taxes from the citizens, residents, and even tourists who stay for enough days a year to qualify. The IRS requires its citizens to file even if they don't live in the country (if you earn above \$9,750) and potentially pay taxes (if you earn above the \$97,600 threshold); the only other country with this overseas form of taxation is Eritrea. Even if you renounce your citizenship, as 3,805 did in 2011, you still have to pay an exit tax of 15 percent on all your assets including investments, homes, and even your personal possessions, literally the shirt on your back.

Extensive data collection

To keep track of this, the IRS has one of the most extensive data collections in the world. Traditionally its power to enforce has come through the matching of data. A W2 (showing how much you earned) submitted to the IRS from your employer should match with what you show on your filing (the deceptively simple-looking 1040 form in the US); 1099 forms show your earning from miscellaneous income, interest, deductions, and so on.

The IRS is, however, not a legacy environment. In many ways it pioneered Big Transaction Processing Data, a task it seems destined to continue with its signing of a \$650 million ten-year contract with Unisys for services and computing, including Unisys's ClearPath Dorado Server solutions for processing tax returns, running at an estimated 1,200 MIPS. These servers will reside in the IRS's Memphis and Martinsburg, West Virginia, datacenters alongside several IBM z/196 mainframes, capable of running at an estimated 8,000 MIPS. Along with all this processing power are the data storage capabilities Unisys has been awarded in this contract, potentially worth \$139 million over ten years. Unisys will transition the management of the IRS's data, estimated to be in excess of 7.5 PB, to a private cloud-based storage-as-a-service solution, providing saleable increases as required for the government agency.



Tax evidence from social

Not satisfied with being the Big Data TP King, the IRS is now going after a new title, the Big Data Social Media Analytics King, but it is currently in research or "stealth" mode. According to information revealed to the American Civil Liberties Union through a Freedom of Information Request, the IRS may currently be reading taxpayers' emails without a warrant. [The ACLU points out:](#)

The federal law that governs law enforcement access to emails, the Electronic Communications Privacy Act (ECPA), is hopelessly outdated. It draws a distinction between email that is stored on an email provider's server for 180 days or less, and email that is older or has been opened. The former requires a warrant, the latter does not.

The position is defended by the IRS. As its Chief Council puts it: "There is no expectation of privacy in electronic communication." Similarly, according to the Electronic Frontier Foundation, in addition to emails, "IRS agents can scan public postings on Facebook as part of research to assist in resolving a taxpayer case," as well as "set up fake social media profiles to attract information and "trick" users into accepting [government] officials as friends."

Unlike the normal corporate big data analytics, the IRS has one big advantage: It knows everyone's social security numbers, as well as all the tax information from the firms we as taxpayers interact with,

MORE BLOGS FROM ROBERT PLANT

A Review: Big Data in 2013

39 comments
 Data scientists, NSA, C-suite buy in... Big data changed a lot in the last 12 months.

What Would Seymour Cray Think of Big Data?

13 comments
 How would the tech legend respond to the 4 Vs?

The 360 Degree View of a Customer

56 comments
 Big data is helping companies like Netflix see what you really want (and even when you go to sleep).

Lessons for Big Data From President Obama's Healthcare Implementation

24 comments
 Three big data lessons for CEOs from HealthCare.gov's bumpy start.

Is Big Data Ruining Sports?

29 comments
 From sailing to the NBA -- big data could increase the gap between sporting haves and have-nots (while ruining the magic in the process).

[More from Robert Plant](#)



FLASH POLL

[All Polls](#)

DIGITAL AUDIO



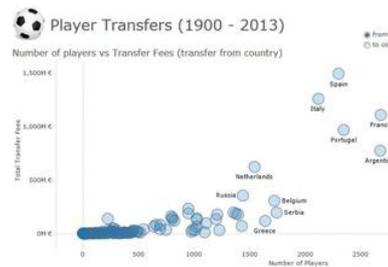
Big Data Analytics Radio

Latest Archived Broadcast
Your Big Data Initiative Starts With a Great Team
 The Big Data Conference and Nemertes Research have teamed up to provide expert advice on building a big data team and finding the right leader. Join us on Tuesday, Aug. 20.
[Listen](#)

[Full schedule](#) [Archived shows](#) [Download to calendar](#)

DATA VISUALIZATION SHOWCASE

Visualizing Sporting Value: Football Player Transfers



A great visualization of transfer fees leads us to think of deeper sporting analytics. Explore this data [here](#).

[More Data Visualization Showcase](#)

and as such can join the dots between Twitter, Facebook, Gmail, and perhaps your credit card along with your emails to overseas bankers. All this will allow it to refine its algorithms such as the "Discriminant Function System." The "DIF" score is one metric the IRS uses to identify candidates for an audit. It compares metrics such as income, deductions, credits, and exemptions for a return against a peer group; when variance is found, an audit may be flagged.

So, while none of us enjoys doing our taxes and complain that our pet tax-funded government initiatives such as a return to the moon may be on hold, we as taxpayers can be comforted by knowing that the government is at the forefront of the big data revolution. Who knows, perhaps it will use these new technology skills to make the government itself more efficient in its other areas and perhaps one day abolish taxes altogether... but I wouldn't count on that any time soon!

Related posts:

- [Payment on \(Data\) Collection](#)
- [Big Data: The Tax Collector's New Friend](#)
- [Big Data to Take the Fraud out of Obamacare](#)

— Robert Plant, Associate Professor, School of Business Administration, University of Miami

[Email This](#) [Print](#) [Comment](#)

COMMENTS

[Newest First](#) | [Oldest First](#) | [Threaded View](#)

PAGE 1 / 2 > >>



Edwin Willems, User Rank: Exabyte Executive
4/30/2013 | 8:04:26 PM

Re: IRS

@robert - although I don't like Big Brother watching all of us, I think tax authorities is a great case to use analytics to leverage a wealth of information, and yes in most cases, people don't know how exposed they are in posting their whole lives.



50% 50%

I can highly recommend this small 'infomercial' to warn people against posting confidential information on the web - watch it!

http://www.rtbf.be/video/detail_une-campagne-de-febelfin-pour-lutter-contre-la-fraude-sur-internet?id=1762087

[Reply](#) | [Post Message](#) | [Edit/Delete](#) | [Messages List](#) | [Start a Board](#)



Susan Fournané, User Rank: Blogger
4/21/2013 | 7:06:12 PM

Re: IRS

Saul,

I see what you mean now. Thanks.



50% 50%

-Susan

[Reply](#) | [Post Message](#) | [Edit/Delete](#) | [Messages List](#) | [Start a Board](#)



Saul Sherry, User Rank: Blogger
4/19/2013 | 7:57:46 AM

Re: IRS

Spot on legalcio, as much as we want to hide it from the officials, once a human has wealth, their first instinct is to flaunt it... our own worst enemy.



50% 50%

[Reply](#) | [Post Message](#) | [Edit/Delete](#) | [Messages List](#) | [Start a Board](#)



Saul Sherry, User Rank: Blogger
4/19/2013 | 7:56:25 AM

Re: IRS

I look at it this way Susan - taking a UK example, we know a load of companies, big big companies, are avoiding paying their tax bills. Mostly it's legal, but even if the government could do something about it, they would be reluctant to do so for fear of driving businesses away. The same goes for individuals, we resisted the banker bonus 'Robin Hood' tax because we were afraid we wouldn't be able to attract the best banking talent.



50% 50%

So who's left as a viable target? Joe Muggs, a little guy with maybe a few minor tax issues. Multiply Joe by a couple of million, and there are untold riches we can find and pump back into the tax system. I'm not saying it's right, but I would imagine social will be able to surface more of the 'everyman' cases.

[Reply](#) | [Post Message](#) | [Edit/Delete](#) | [Messages List](#) | [Start a Board](#)



legalcio, User Rank: Exabyte Executive
4/18/2013 | 4:49:58 PM

Re: IRS

Crooks are, by and large, stupid. They are especially stupid when it comes to social media. They were pretty stupid even before social media when all the Feds had in terms of tools were wiretaps. If John Gotti had just kept his mouth shut he wouldn't have died in prison.



50% 50%

I suspect the IRS will go after the low hanging fruit in social media, triggered by a profile that boasts of, or infers high income and matching it with zero taxes paid.

Want to stay off the government radar? Don't use social networks and don't have a Social Security number.

[Reply](#) | [Post Message](#) | [Edit/Delete](#) | [Messages List](#) | [Start a Board](#)



Susan Fournané, User Rank: Blogger
4/18/2013 | 12:43:22 PM

Re: IRS

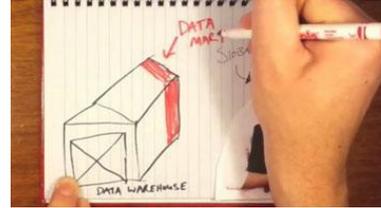
Saul,

BDR IN YOUR INBOX

Enter email

FEATURED VIDEO

Big Data Explained: What is a Data Mart? 18



Use this video to explain quickly to someone the difference between a data warehouse and a data mart. [Watch This Video](#)

[More Video Blogs](#)

FOLLOW US ON TWITTER

Big Data Republic @BigDataRepublic 15 Apr
AllJoyn: A Common Language For Internet Of Things - Open-source lingua franca aims to let IoT devices and services... [ow.ly/2FHbY7](#)
Retweeted by Yakamoto
Expand

Hajime Ozaki @hajozaki 22h
BDR Top 20 #BigData David Smith David Smith List of top tweeters on Big Data from @BigDataRepublic [wp.me/p4Aavi-2](#)
Show Summary

Marcus Borba @marcusborba 28 Apr
#BigData #Analytics #BusinessIntelligence #FS @informationweek @AllAnalytics @BigDataRepublic @BigDataNetwork @BigDataClub

LIKE US ON FACEBOOK

Big Data Republic
Like

4,168 people like Big Data Republic.

Facebook social plugin

ACCOLADES



[Login](#)



Well, yes. Going after the elusive makes sense, of course. Then again, to what extent the obvious are hiding from the IRS with so much exposure everywhere?

But the little numbers, those who don't even have a little car, (although little cars don't exist in the U.S. :) is there a point for the IRS to check the social media of those?

-Susan

[Reply](#) | [Post Message](#) | [Edit/Delete](#) | [Messages List](#) | [Start a Board](#)



Saul Sherry, User Rank: Blogger
4/18/2013 | 12:05:59 PM

Re: IRS

That said, a lot of little numbers would add up to some hefty income - machine learning, image recognition and refinement of the coding which links all three could yield some huge results here. I can't imagine they will be going into this field to chase the obvious, they want the elusive.

[Login](#)



[Reply](#) | [Post Message](#) | [Edit/Delete](#) | [Messages List](#) | [Start a Board](#)



Saul Sherry, User Rank: Blogger
4/18/2013 | 12:04:20 PM

Re: IRS

And the photos are probably pretty compelling evidence for inspectors - that's before we take into account Geo-location...

[Login](#)



[Reply](#) | [Post Message](#) | [Edit/Delete](#) | [Messages List](#) | [Start a Board](#)



James M. Connolly, User Rank: Blogger
4/18/2013 | 11:31:07 AM

Re: Great information!

Saul. It's a pretty safe bet that more nefarious organizations are already milking social media outlets for details about our personal lives and businesses. At minimum they are using it to target people with social engineering. At worse they are using it for crimes like extortion. The latter is why security experts advise prominent CEOs not only to limit their own social media activity but to have their spouses and children limit their own activity to prevent evil-doers from patterning their lives and travels.

[Login](#)



[Reply](#) | [Post Message](#) | [Edit/Delete](#) | [Messages List](#) | [Start a Board](#)



Susan Fourtané, User Rank: Blogger
4/18/2013 | 10:50:04 AM

Re: IRS

Ariella,

Yes, that's true. I guess they will not waste time and resources in checking the social media of those who don't represent big numbers. What would be the point?

[Login](#)



-Susan

[Reply](#) | [Post Message](#) | [Edit/Delete](#) | [Messages List](#) | [Start a Board](#)

PAGE 1 / 2 > >>

BIG DATA REPUBLIC

[ABOUT US](#) | [CONTACT US](#) | [HELP](#) | [REGISTER](#)
[TWITTER](#) | [FACEBOOK](#) | [LINKED IN](#) | [GOOGLE+](#) | [RSS](#)



bigdatarepublic : /bigdatarepublic/section/2635 : /bigdatarepublic/section/2635/262154



UBM TECH

OUR MARKETS: **Business Technology** | **Electronics** | **Game & App Development**

Working With Us: [Advertising Contacts](#) | [Event Calendar](#) | [Tech Marketing Solutions](#) | [Corporate Site](#) | [Contact Us](#) / [Feedback](#)

[Terms of Service](#) | [Privacy Statement](#) | [Copyright © 2014 UBM Tech, All rights reserved](#)