



# *Torturing Open Government Systems for Fun, Profit and Time Travel*

Dr. Tom Keenan FCIPS, I.S.P., ITCP, Professor  
Faculty of Environmental Design & Dept. of Computer Science

[keenan@ucalgary.ca](mailto:keenan@ucalgary.ca)

[@drfuture](#)

[decreeping.wordpress.com](http://decreeping.wordpress.com)

**DEF CON 21**

**Las Vegas, NV**

**August 1-4, 2013**



# Disclaimers

- These views are mine, not my employer's, my publisher's, my cat's etc.
- I DO LIKE OPEN DATA! I just want to see it done right.
- Some of the vulnerabilities shown have been fixed/changed/modified. They are still instructive as cautionary tales.
- More results will appear in a forthcoming book from OR Books and on my blog [decreeping.wordpress.com](http://decreeping.wordpress.com)

# The Six Ws of Open Data

- What is Open Data?
- Who has Open Data?
- Why is Open Data?
- Where is Open Data?
- When is Open Data?
- Worrisome aspects of Open Data

# What is the Open Data Movement?

- Initiative by many governments to make their data openly available
- Driven by:
  - Calls for transparency (“it’s the taxpayers money”)
  - Sincere desire for innovation (“they’ll come up with some great uses for our data at that Hackathon”)
  - Fear (“Wikileaks is going to get this anyway; it looks better if we release it”)
  - Cheapness (“we don’t have to pay the developers”)
  - Positive sentiments about open source, transparency, “data wants to be free”



# Some notable examples

- NYC Datamine
- Philadelphia OpenDataPhilly
- Open Data Framework (Vancouver, Edmonton, Toronto and Ottawa)
- Open Data Challenge (EU)
- Direct input systems like [seeclickfix.com](http://seeclickfix.com)

# NYC Datamine

- On October 6, 2009 NYC released, with great fanfare, 103 municipal data sets
- One listed 1100 women's organizations in the city
- Release contained their private email address, secret question like "favorite pet" and answer like "fluffy"

**Problem #1: Neglecting to read and redact data sets before releasing!**

# Toronto's 311 Customer Requests System

- 25% are about garbage/recycling pickup
- Some are serious, personal etc.
- Location is *supposed* to be reported as partial postal code e.g. M4Vxxx, but e.g. WOLFE AVE & DANFORTH RD
- If there are enough of these slip ups, combined with Google maps, etc. they can tell a story...

Problem #2: Sloppy, lazy data entry

# See. Click. Fix. Shame.



## Large bump/hole in the road • Open

Queen Elizabeth II Hwy Calgary, AB T3N 1C5, Canada - Calgary

On the right hand lane, just as the bend begins to go right onto the Deerfoot, on the left side of the lane there is a huge dip/bubble/hole (hard to describe it) that if you hit it, can cause serious issues. Motorcyclists can actually 'get air' if they hit it dead-on and that could lead to... [read more](#)

2  
votes

FIX IT!

2011-05-17 · Reported by Ed Reddy (Guest) · Share · Flag



Ed Reddy

The bump on the road is getting worse, and now it looks like it has a friend on the other side of the lane. It will require fixing shortly unless there's an accident first.

2011-11-02 · Flag

Write a comment...



## Excessive Car Wash Noise • Open

116 89 St Sw Calgary, AB - Calgary

Car wash produces daytime noise exceeding legal limit of 65 dBA. Finally banned from operating at night time but often still violates by opening early on Sundays and Stat Holidays.

1  
vote

FIX IT!

2013-01-21 · Reported by Buster (Guest) · Share · Flag

Write a comment...



## Dangerous Ice on Road from Car Wash • Open

116 89 St Sw Calgary, AB - Calgary

Dangerous ice accumulating on Old Banff Coach Road due to dryer problem from Petro-Canada Glide

1  
vote

FIX IT!

2012-12-08 · Reported by Buster (Guest) · Share · Flag

Write a comment...





# When you offer a 20,000 € prize, things happen!



The Open Data Challenge is over!



- What was it?
- What did people say about it?
- Who were the winners?
- Who were the judges?
- Who was behind it?
- Where can I find open data?

If you'd like to keep in touch with other people interested in **open data**, you can join the **open-government** and **euopendata** mailing lists or follow the **#opendata** hashtag on Twitter.

# Who Won?



PublicData.eu **beta** — Europe's Public Data

Groups Apps Ideas About

Login Register

Find datasets

## ZNasichDani / From Our Taxes

View

### Link

<http://znasichdani.sk/?l=en>

### Developer

Fair-play Alliance and VORSR.sk

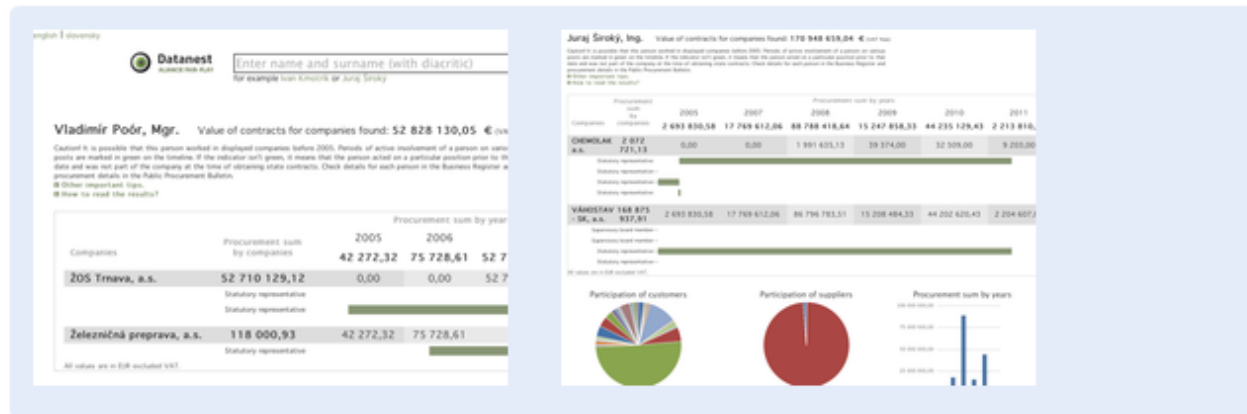
### License

Closed

### Tags

power networks transparency managers politicians

ZNasichDani.sk uncovers who are influential persons (owners, managers, statutes) standing behind companies successful in securing contracts with the state, thus helping uncover conflicts of interest; it also shows how large are the sums of contracts obtained in this way.



Source: [www.fair-play.sk](http://www.fair-play.sk)

# Who Wouldn't Agree with Such a Wonderful Application?

## Open Data Challenge Winner Ordered To Remove Certain Data

Source: [spectator.sme.sk](http://spectator.sme.sk)

A court has ordered Slovak NGO Fair-Play Alliance to take down data from their award winning open data application.

Bratislava, 28 June 2011

(by Ton Zijlstra)

The Slovak NGO [Fair-Play Alliance](#), a transparency watch dog organisation, won the [Open Data Challenge](#) with their application [Znasichdani.sk](#) during the Digital Agenda Assembly in Brussels on June 16th.

A Slovak judge has now ordered the NGO to take down specific data from their winning application.

The application combines two public registers. The overview of all work contracted by the state, and the company register are combined to show what the total volume of work is that various people, through the businesses they represent or own, do for public bodies.

According to a [news article in the Slovak Spectator](#) the statutory representative, Jarmila Povazanova, of the Strabag construction company demanded in court that the total value of all public contracts of companies Povazanova represents be removed from the site. The Bratislava II district court ordered the NGO to remove that information. Director Zuzana


# A little technical aside

- This thing runs off multiple open databases all maintained by the Slovak government
- The Fair-Play Alliance has read only access
- How are they supposed to “take down specific data” that they don’t control?
- Fair-Play Alliance has appealed the ruling and is still in operation.

Problem #4: Laws, policies and judges lag behind technology



# Of Course, Election Results Should be Open Data (right?)



The screenshot shows the City of Edmonton Open Data Catalogue interface. At the top, there's a header with the city logo and navigation tabs for 'CITY ADMINISTRATION', 'DEMOGRAPHICS', and 'EDUCATION'. Below the header is a breadcrumb trail: 'Home > City Government > OpenGov > Open Data Catalogue'. The main content area shows an 'Unsaved View' for 'Based on 2010 Election Mayor Voting Summary' with 'Final 2010 election results for various municipal and school board positions'. A table displays the results with columns for 'CandidateName' and 'VotesReceived'. The table has 9 rows, with the second row highlighted in blue. A red box highlights the 'VotesReceived' column for the fifth candidate, Dave Dowling.

	CandidateName	VotesReceived
1	Stephen Mandel	109,432
2	David Dorward	58,856
3	Daryl Bonar	20,672
4	Dan Dromarsky	4,621
5	Dave Dowling	2,295
6	Under Votes	1,141
7	Andrew Lineker	1,131
8	Bob Ligertwood	1,129
9	Over Votes	82

# Did My Wife Vote for Me?

City	Ward 11 City Wide Hospital	9	FALSE	1	Daryl Bonar	1
City	Ward 11 City Wide Hospital	9	FALSE	2	David Dorward	5
City	Ward 11 City Wide Hospital	9	FALSE	3	Dave Dowling	0
City	Ward 11 City Wide Hospital	9	FALSE	4	Dan Dromarsky	0
City	Ward 11 City Wide Hospital	9	FALSE	5	Bob Ligertwood	0

Problem #5: Inference from Small Numbers

# Philadelphia's OpenDataPhilly.org

- Includes Campaign Finance Records, which are generally required to be disclosed.
- “Results are freely viewable and printable but not available for download” and “empty searches take too long”
- It wasn't long until I had me some .csv files to play with

# What We Can Learn

- Full “home” address provided. Tied to tax receipt so probably chosen with care
- I’m not sure what percentage of people try to protect their address, but in the US ‘Nationally, 27.6% of households have unlisted (telephone) numbers. In many cities the unlisted rate is over 50%.’

- Source: <http://www.busreslab.com/index.php/articles-and-stories/research-tips/general-research-tips/the-science-of-sampling-telephone-samples/>

- People often pay for privacy, need it for personal for business, reasons, etc.



# Campaign Finance Disclosure Rules Date Back to 1970s

- US Federal Elections Commission files go back to 1980. Just a reminder of what we had for data mining technology back then:




Source: Wikimedia Commons

# There are some rather interesting addresses in the file

SHAWN KIRWIN	1719 SPRING GARDEN STREET Philadelphia, PA 19130-3915	\$350.50	Jan 25, 2013	International Brotherhood of Electrical Workers Local 98 Committee on Political Education
HABIB WALEED	4959 N FRANKLIN STREET PHILADELPHIA, PA 19123-1326	\$5.00	Apr 15, 2011	International Union of Painters and Allied Trades Political Action Together Political Committee
PAUL E PROIETTO	691 MILLBROOK RD. RIVER EDGE, NJ 07661-1453	\$16.36	Feb 25, 2011	International Union of Painters and Allied Trades Political Action Together Political Committee
Ronald Rivest	41 Academy Street Arlington, MA 02476	\$200.00	May 4, 2011	Friends of Stephanie Singer

# Hmm, is that THE Ron Rivest?



**FEDERAL ELECTION COMMISSION**

[About the FEC](#) [Press Office](#) [Quick Answers](#) [Contact Us](#) [Site Map](#)

[HOME](#) / [DISCLOSURE PORTAL HOME](#) / [NATIONAL MAP](#) / [TRANSACTIONS](#)

## 2012 Presidential Campaign Finance

**Current Search Criteria:** Contributor's Name Begins with "RIVEST" for All Candidates

[Modify Search](#) [New Search](#)

Export Options: [XML](#) [CSV](#)

Prev **1** Next [Show All](#) | Total Result(s):4

Contributor Name (Contribution Description)	Employer/Occupation	City	State	Zip	Receipt Date	Amount	Candidate Name
RIVEST, ALEXANDER	MIT/POST-DOC	ARLINGTON	MA	02476	10/22/2012	\$250	Barack Obama
RIVEST, GAIL D.	NOT EMPLOYED/SCHOOL VOLUNTEER	ARLINGTON	MA	02476	10/23/2012	\$2,500	Barack Obama
RIVEST, RONALD L.	MIT/PROFESSOR	ARLINGTON	MA	02476	08/05/2012	\$2,500	Barack Obama
RIVEST, RONALD L.	MIT/PROFESSOR	ARLINGTON	MA	02476	11/01/2012	\$1,000	Barack Obama

Prev **1** Next Total Result(s):4

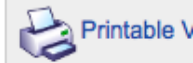
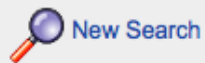
- Data Catalog
- Maps
- Presidential
  - House & Senate
  - House Independent Expenditures
  - Senate Independent Expenditures
- Charts
- Search
- Download

# Search for Smith in 2010 Political Contribution Records



phila.gov | records department

[Document Search](#) | [Logout public](#)



**You searched for:** Year = 2010 and approximate search in ContributorID for Smith and =CFR - Schedule I - Part A - Contributions Received From Political Committees (\$50.01 - \$250.00), CFR - Schedule I - Part B - All Other Contributions (\$50.01 - \$250.00), CFR - Schedule I - Part C - Contributions Received From Political Committees (Over \$250.00), Schedule I - Part D - All Other Contributions (Over \$250.00), CFR - Schedule I - Part E - Other Receipts, CFR - Schedule II - Part F - In-Kind Contributions Received (Value of \$50.01 - \$250.00), CFR - Schedule II - Part G - In-Kind Contributions Received (Value Over \$250.00)

[Download Results as CSV](#)

Contribution By	Address	Amount	Date	Recipient
D. Smith	Two Liberty Place 1601 Chestnut St Philadelphia, PA 19192	\$6.83	May 13, 2010	CIGNA Corporation Political Action Committee
Joseph Smith	1719 Spring Garden Street Philadelphia, PA 19130	\$101.20	Nov 25, 2009	Local Union #98 IBEW COPE
D SMITH	1534 E 80TH ST LOS ANGELES, CA 90001-3404	\$40.00	Apr 22, 2010	Communication Workers of America
LETITIA SMITH	16368 KENSINGTON PLACE MORENO VALLEY, CA 92551-0000	\$10.00	Apr 26, 2010	Communication Workers of America
Mark Smith	Two Liberty Place 1601 Chestnut St Philadelphia, PA	\$5.00	Aug 19,	CIGNA Corporation Political Action Committee



# Why Include Address?

## Laziness? Disambiguation?

- The donor files for 2010 were downloaded from OpenDataPhilly as .CSV files, imported into Microsoft Excel and sorted by the contributor name field.
- “Potential Duplicates” were defined as additional names that were shown identically as a name in the database, but with a different address.

# Results from most common US names

Name	Number of Entries	“Dups”
Smith	588	8
Johnson	426	3
Williams	400	4

Conclusion: Having home address didn't do a lot of good in resolving possible duplicates.

# Are those addresses real home addresses?

- A significant number of contributors (86/588 of Smith, 63/426 of Johnson, 77/400 of Williams) listed the same address, 1719 Spring Garden Street, Philadelphia, PA., 19130.
- Viewing this location on Google Maps Streetview shows that it is the office of the I.B.E.W. Electricians Union.





Home

Who We Are

Why Union?

Education

Links

Contact Us

Departments

Guest Book

Site Map

## Departments - Contact Us

To contact us via Email, please click [here](#).

### IBEW Local 98 Business Office All Divisions

1701 Spring Garden Street  
Philadelphia, PA 19130

Phone: 215-563-5592  
Fax: 215-561-2168



### Financial Office & Electrical Apprentice Training

1719 Spring Garden Street  
Philadelphia, PA 19130

Financial Office  
Phone: 215-563-8991  
Fax: 215-563-9149

Apprentice Training  
Phone: 215-567-6405  
Fax: 215-567-6477





# Problemas ?

- #6 Under-estimating the desire and ability of someone to torture the database for info
- #7 Inconsistent data. Providing contributor's "address" is a legal requirement but some people use non-home, e.g. union office
- #8 The "why" e.g. of requiring "address" needs to be re-examined in the light of new "hows" -- tools like mapping APIs, big data, jigsawing of public and private databases

This means we'll be flying to different cities because they'll give  
y to tie into their networks. It will be both international carriers  
ada and transporter carriers that we'll be able to feed into.  
e of an exciting year in the sense that all of our old airplanes  
2000) are gone, and we have the newest fleet in Canada—in fact

This is an old (6 years ago) jigsawing experiment I did...  
Some things have gotten better, many have gotten worse!

yourself  
Clive Beddoe,  
then-CEO of  
WestJet  
Airlines



**What do you like to  
on WestJet's  
live seatback TV?**

I watch the hockey g  
There's also a busin  
The Report on Busines  
there's a hockey gam  
certainly nice to be a  
watch the Flames pl  
not a great TV watch  
any of the education  
programs, like the D  
Channel, are always  
interesting.

**What is the key to  
WestJet's succes**

It's our commitme  
people, and they  
have a similar cor  
to us. It's about a  
the interests of o  
with the compar  
interests of the c  
aligning with the  
people. And in t  
created a trem  
powerful custo  
service culture

ently  
to  
c.  
nts on  
to

No,  
at sort

Search: "beddoe, AB" 6 Total Listings

[New search](#) | [Modify search](#)

**BEDDOE, C** 1 of 6  
Calgary, AB T1X 1E1 [More information](#)  
(403) 255-2686

[Check out C Beddoe with Web Detective!](#)  
[Search School Yearbooks in Calgary AB](#)  
[Find Romance in Calgary at Date.ca](#)

**BEDDOE, CLIVE** 2 of 6  
1... Ave SW [More information](#)  
Calgary, AB T2V 0R2  
(403) 255-...

[Check out Clive Beddoe with Web Detective!](#)  
[Search School Yearbooks in Calgary AB](#)  
[Find Romance in Calgary at Date.ca](#)

**BEDDOE, J** 3 of 6  
Calgary, AB T1X 1E1 [More information](#)  
(403) 521-2260

[Check out J Beddoe with Web Detective!](#)  
[Search School Yearbooks in Calgary AB](#)  
[Find Romance in Calgary at Date.ca](#)

**BEDDOE, K & S** 4 of 6  
Calgary, AB T1X 1E1 [More information](#)  
(403) 212-2352

[Check out K & S Beddoe with Web Detective!](#)  
[Search School Yearbooks in Calgary AB](#)  
[Find Romance in Calgary at Date.ca](#)

**BEDDOE, O** 5 of 6

censored...phone number is still valid!



# City of Calgary Property Tax Assessment Site in 2007



- Home
- City Business
- City Hall
- City Living
- City Transportation
- My Pages

City Hall > Business Units > Assessment

## Assessment

- Assessment
- About Assessment
- 2006 Assessment Values - Quick Facts
- Related Contacts
- Contact Us
- Property Assessment
- Business Assessment
- Understanding your Assessment Notice
- Market Value
- 2006 Assessment Values
- Community Comparisons & Profiles

## Assessment Search

You may search all residential, multi-residential, and non-residential property types in The City of C

SEARCH BY  
MAP

Works best using cable modem or (A)DSL and using Microsoft IE or Netscape versions 4 and up.

SEARCH BY  
ADDRESS

Works best with Internet connection speed better and using Microsoft IE or Netscape up.

**You must know the exact address of property. For condominium properties building address and a list of units w provided.**

Assessment Information Last Updated: April 19, 2006

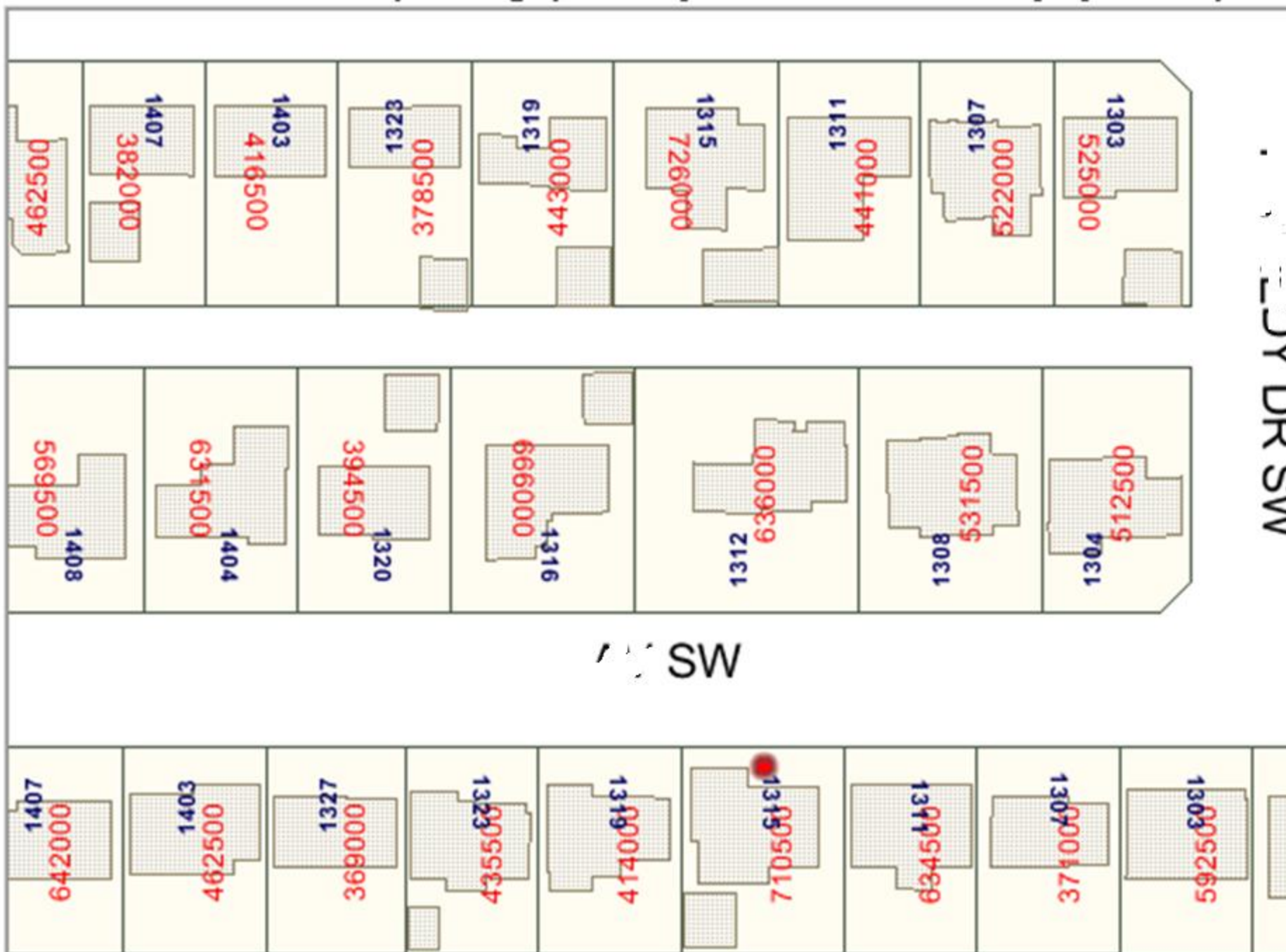




## Print Neighbourhood Map

Your property assessment is based on The City of Calgary's estimate of the market value December 31, 2005.

<b>Roll number:</b>	111
<b>Property Location:</b>	11111 AV SW
<b>Property Assessment:</b>	710,500
<b>Assessment Class:</b>	Residential
<b>Property Type:</b>	Land and Improvement
<b>Property Use:</b>	Single Residential
<b>Taxation Status:</b>	Taxable
<b>Community Name:</b>	Kelvin Grove
<b>Land Use Designation:</b>	RESIDENTIAL SINGLE-DETACHED DISTRICT
<b>Parcel Area:</b>	8,095 sq. ft. / 752 sq. m.
<b>Building Description:</b>	House/2 Storey Garage
<b>Approx. Living Area above grade:</b>	3,260 sq. ft. / 303 sq. m.
<b>Finished Area below grade:</b>	Yes
<b>Garage:</b>	Yes
<b>Fireplace(s):</b>	Yes
<b>Walkout Basement:</b>	No

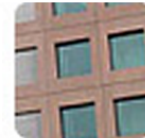


# Now “addressed” by a combination of data changes, technical and legal safeguards



THE CITY OF  
CALGARY

## Terms of Use



Follow us on



**ACCEPTANCE OF TERMS** Please review these Terms of Use carefully. If you access, use or purchase a product or service on a Web site or Web page ("Site") provided by The City of Calgary ("The City") you accept and agree to comply with the following Terms of Use. If you do not agree with these Terms of Use, you should immediately discontinue any use of this Site. These Terms of Use may change from time to time. You should visit this web page periodically to review the Terms of Use that govern your transactions and are binding on you.

**LICENSE** The City grants you a limited non-exclusive, non-transferable license to access and make personal use of the Site for non-commercial use, provided the Site is not modified or downloaded (other than page caching) unless expressly authorized by The City. You will be fully responsible for any consequences resulting from any use. **This license does not include any right for resale or commercial use of the Site or its content. No**

**I Accept The Terms**

**I Do Not Accept The Terms**

# Don't get arrested in Florida



## Hernando County Sheriff's Office

Sheriff Al Nienhuis



- HCSO Home
- Explore HCSO
- Hernando County Jail
- Emergency Management
- Traffic Information
- Alert Bulletins
- Crimestoppers
- Media Releases
- CASE Partnership
- Fraud Alerts
- General Information
- Public Records Search
- Crime Statistics
- Sexual Offender Info
- Job Line
- Alert Notification

### About Hernando County Sheriff's Office

Hernando County is located on the West coast of Central Florida. The Sheriff's Office provides law enforcement services to a population of approximately 173,000. This is an increase of nearly 330% since 1980, and Hernando continues to be regarded as one of the fastest growing counties in the United States.






This decade will bring new challenges to our county, but the commitment of the Hernando County Sheriff's Office will not change. We will continue to provide our community with professional, efficient law enforcement through well-trained employees, technology, crime prevention and a proactive approach to public safety.



### Highlights

- [Click here for Info on how you can purchase t-shirts from the Mecklenburg Memorial Motorcycle Ride](#)
- [Run for New Beginnings Youth Shelter 5K Run/Walk](#)
- [Click here for more information about the Sheriff's Office proposed BUDGET for 2012](#)
- [Public Notice - Edward Byrne Memorial Justice Assistance Grant](#)
- [HCSO Citizen Volunteer Corps](#)
- [Pictures from 2011 Law Enforcement Special Olympics Torch Run](#)
- [Hernando County Community](#)



Report#:	ArrestDate:	Name:	DOB:	Offense:	Arrest Location:	Latest Photo Available:
2009-04928/1	02/06/2009 11:44 PM	SIMMONS, CATHY L	02/19/1961	831.02 - UTTERING FORGED INSTRUMENT	6242 CALIFORNIA ST	
2009-04923/1	02/06/2009 11:05 PM	WILBUR, WILLIAM LEO III	01/26/1941	88.231 - CONTEMPT COURT/CHILD SUPPORT	16160 APPENNINE DR	
2009-04913/1	02/06/2009 09:41 PM	LEDE, JEANETTE MICHELE	10/01/1981	322.34 - DRIVER'S WITH LICENSE REVOKED OR SUSPENDED	16425 SPRING HILL DR	
2009-04912/1	02/06/2009 09:41 PM	LEDE, JEANETTE MICHELE	10/01/1981	948.06 - VIOLATION OF PROBATION	16425 SPRING HILL DR	
2009-04909/1	02/06/2009 08:55 PM	ASKEW, ROBERT JARMAN	10/31/1959	784.03 - DOMESTIC BATTERY	16096 MEDRICK RD	

# What if this was Your Son?

- 01/08/2010 04:00PM
- COU\*\*\*EY, BOBBY
- DOB 11/10/1997
- (age = 12 yrs. 2 mos)
- 810.02 - BURGLARY DWELLING
- 18900 CORTEZ BV



Problem #10: Your face is your newest privacy risk as facial recognition goes mainstream.

# Data Journalism: Playing with Pistol Permits



Well, we could let it stop there — knowing this is a homeowner you shouldn't mess with.

But, let's face it, we have tools like Spokeo and Maltego. Even without paying a fee, Spokeo confirms the poor fellow's street, gives us his middle initial (E), his approximate age (early 50s) and the value of the property he's defending with that handgun (\$457K). In all probability he's also the same person who owns a post office box in the same zipcode.

Original: FOI request by The Journal News, lohud.com

# Provocative Question: Should the Rich Have Less Privacy?

- People who run companies that get contracts, give contributions (especially large ones) have money
- In the US there is some sentiment that the SNAP Food Stamp program is abused
- In theory we could post that Ms. X bought junk food (allowed) and attempted to buy beer with her food stamps...but we don't.



# Indirect Risks of Government Disclosure

- You can't get it back. Companies like ChoicePoint Systems (now part of Lexis/Nexis) are grabbing and storing it
- New technologies (think DNA matching)
- “Genetic Genealogy”

# Ancestry.com

- 14 years of operation
- 6,000,000,000 records, mostly from public sources
- 20,000,000 family trees

NEW!



## One simple DNA test. A world of discoveries.

Discover if you're part Scandinavian, West African, or maybe Native American. Meet a third cousin for the first time. When combined with 34 million family trees and 10 billion records on Ancestry.com, the new AncestryDNA creates the most comprehensive family history experience yet. **It's family history, reinvented.**

Now \$99

[Get AncestryDNA](#)



[Experience AncestryDNA](#) [Discover your ethnicity](#) [Find new relatives](#) [How it works](#)

It's more than a DNA test.  
It's a dynamic new experience in family history.

# Or try the mobile service



Problem #11: Information about DNA is getting into our hands (e.g. from 23andme.com) and where else?



# Beware of Time Traveling Robots from the Future

# PRIVACY INTERNATIONAL

[DONATE](#)[ABOUT US](#)[RESEARCH](#)[KEY RESOURCES](#)[BLOG](#)[PRESS AND NEWS](#)[EPHR](#)[PRIVASIA](#)[Home](#)

## PI FILES COMPLAINT ABOUT ONLINE DNA GENEALOGICAL TESTING FIRM

OCTOBER 22, 2007 BY *PI*

The leading online genealogical site Ancestry.com has embarked on a new service offering to test the DNA of its customers. This commercial service has been provided to assist genealogical tracing through familial lines.

However there are a number of troubling questions that Privacy International believes must be resolved before people can have confidence in this new service. Until these questions have been satisfactorily answered, we believe the service presents a number of substantial dangers for customers.

<https://www.privacyinternational.org/>

[Search](#)

### MORE LIKE THIS

- [Genewatch campaign to 'Reclaim your DNA' from growing UK database](#)
- [PI Files complaint about expansion of CCTV on Toronto transit network](#)
- [Police are allowed to Keep DNA of Innocent, say UK Law Lords](#)
- [Key report damns UK National DNA Database Policy](#)
- [PI files complaint about Google Street View](#)

# Tools

- An inquiring mind
- A motive (financial, ID theft, journalism, whistle blowing)
- Scripting languages
- ScraperWiki, collection of Ruby, Python, PHP scripts and plenty of ideas
- SQL Lite and other databases
- Hackathons

Problem #12: Data collected now, and in the past, can will be analyzed, cross-referenced, etc. in the future

# Even if it's always been public, now it's "super-public"

- Data that used to be public but buried in a dusty courtroom basement
- Now instantly available, searchable, storable
- "if it happened in the past ten years, it might be online. If it happened in the past five years, it's probably online. And if it happened in the past two years, it's almost certainly online."

(Source: Fertig and Thompson, Wild West 2.0,  
ISBN-13: 978-0-8144-1509-2)

# The Dirty Dozen (and growing)

- Inadvertent disclosure
- Sloppy data entry
- Malicious misinformation
- Law/Policy lag
- Inference from Small Numbers
- Assuming nobody will torture data



# The Dirty Dozen (and growing)

- Inconsistent data
- Ignoring new tools like mapping, visualization
- Data jigsawing
- Facial recognition
- Powerful data like genetic info in hands of the consumer
- Retroactive analysis with better tools, databases, etc.

# Some recommendations

- Scan files carefully for direct PII that may be included;
- Consider ways in which PII may be revealed indirectly;
- Act promptly to remove database that are shown to reveal PII, and retain clear legal rights to do so;
- Anticipate the cross-correlation of government data with other databases, public and private;
- Provide a convenient mechanism for users to express privacy concerns and follow up;
- Sponsor hackathons before the data is released to try to foresee unanticipated uses;
- Negotiate strong privacy protection on data provided to the private sector;

# Thank you!

- Dr. Tom Keenan, FCIPS
- [keenan@ucalgary.ca](mailto:keenan@ucalgary.ca)

This image courtesy  
Elise Pelchat  
[www.conceptsforall.uqam.ca](http://www.conceptsforall.uqam.ca)

