

What is Open Data ?

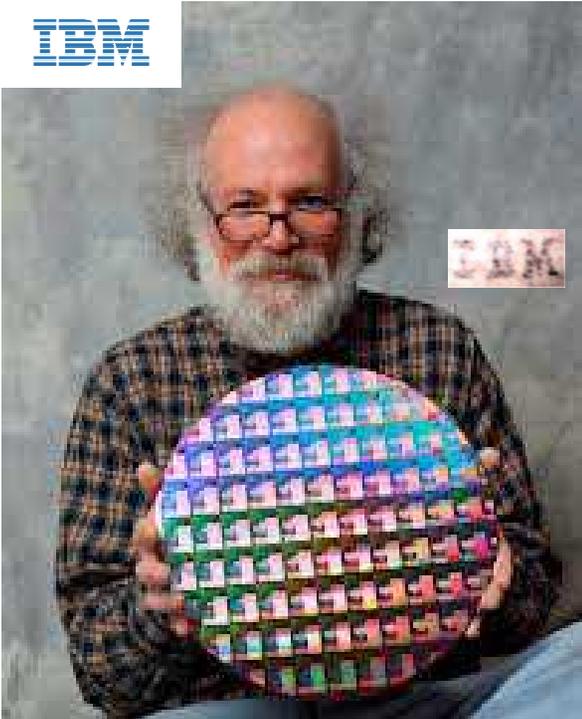
And why should we care ?

Oct 8, 2013



Where I'm coming from on Open Data

IBM



What *is* Open Data ?

Open data is:

- Data that is generally provided at no cost and with no license constraint
- Machine readable and often people readable
- Accessed via files or APIs especially for linked data. The data may be static or dynamic.

Why is Open Data being used ?

- **Transparency** : to build stakeholder understanding and confidence. Increasing data visibility both inside and outside the organization
- **Cost efficiency**: to reduce the cost and effort of developing, operating and delivering services (both IT and civic services)
- **Economic** : to enable innovation by small agile businesses to develop new insight, new services and new products (e.g. tools for 'co-creation', innovative systems of engagement)

Who are using Open Data ?

- Governments agencies and the research community
- Private industry looking for a competitive advantage

Eight principles of open government data*

1. **Data Must Be Complete** - subject to not compromising privacy.
2. **Data Must Be Primary** - data are published as collected at the source, with the finest possible level of granularity, not in aggregate or modified forms.
3. **Data Must Be Timely** - data are made available as quickly as necessary to preserve the value of the data.
4. **Data Must Be Accessible** - data are available to the widest range of users for the widest range of purposes.
5. **Data Must Be Machine processable** - data are reasonably structured to allow automated processing (NOT .pdf)
6. **Access Must Be Non-Discriminatory** - data are available to anyone, with no requirement of registration.
7. **Data Formats Must Be Non-Proprietary** - data are available in a format over which no entity has exclusive control.
8. **Data Must Be License-free** - data are not subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed as governed by other statutes.

* <http://www.opengovdata.org/home/8principles>

Open Data is big and growing

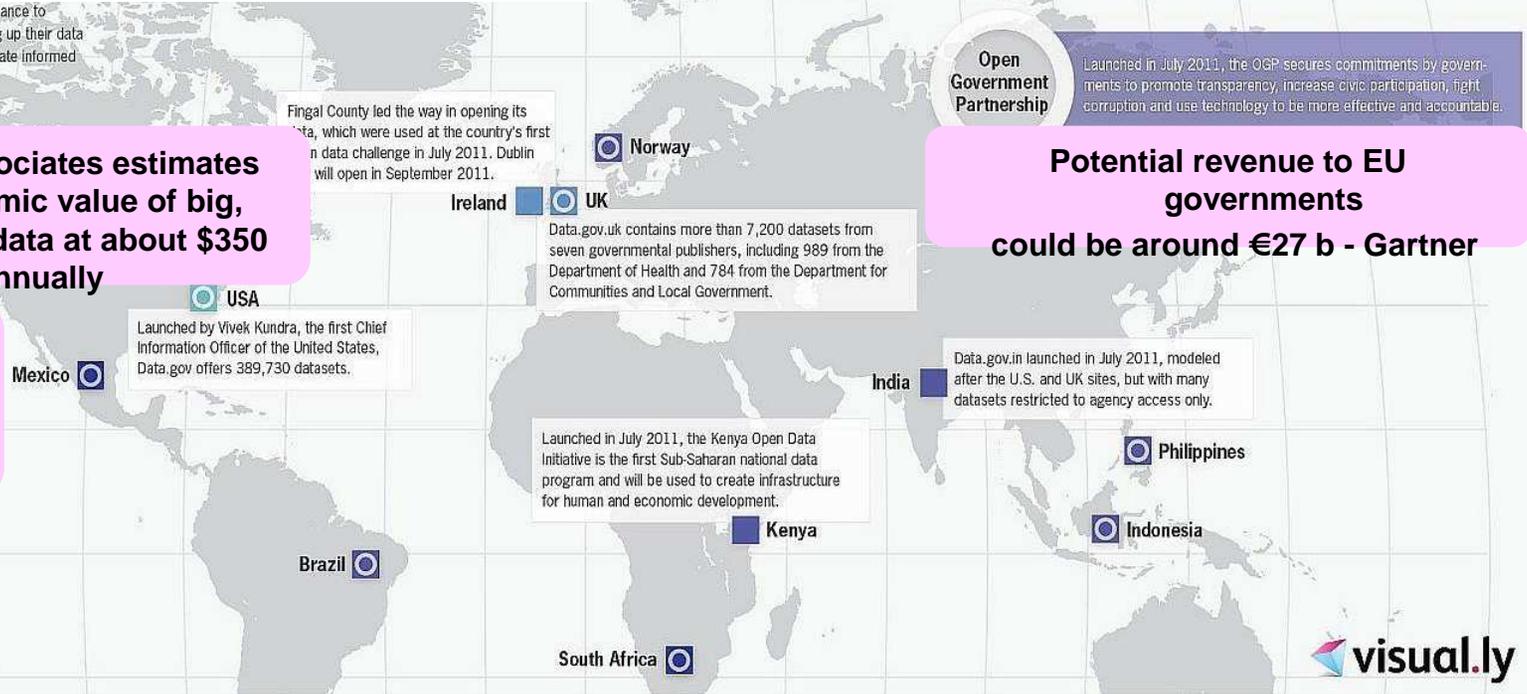
Source: ProgrammableWeb

From education to energy, health to poverty, and finance to demographics, governments and NGOs are opening up their data troves so that anyone can look for patterns and create informed solutions to global challenges.

McKinsey and Associates estimates the annual economic value of big, open liquid health data at about \$350 billion annually

Open data today: 162 governments 43 countries > 1 million data sets

 OGP COMMITTEE MEMBER
 GOVERNMENT DATA RELEASE





Potential revenue to EU governments could be around €27 b - Gartner

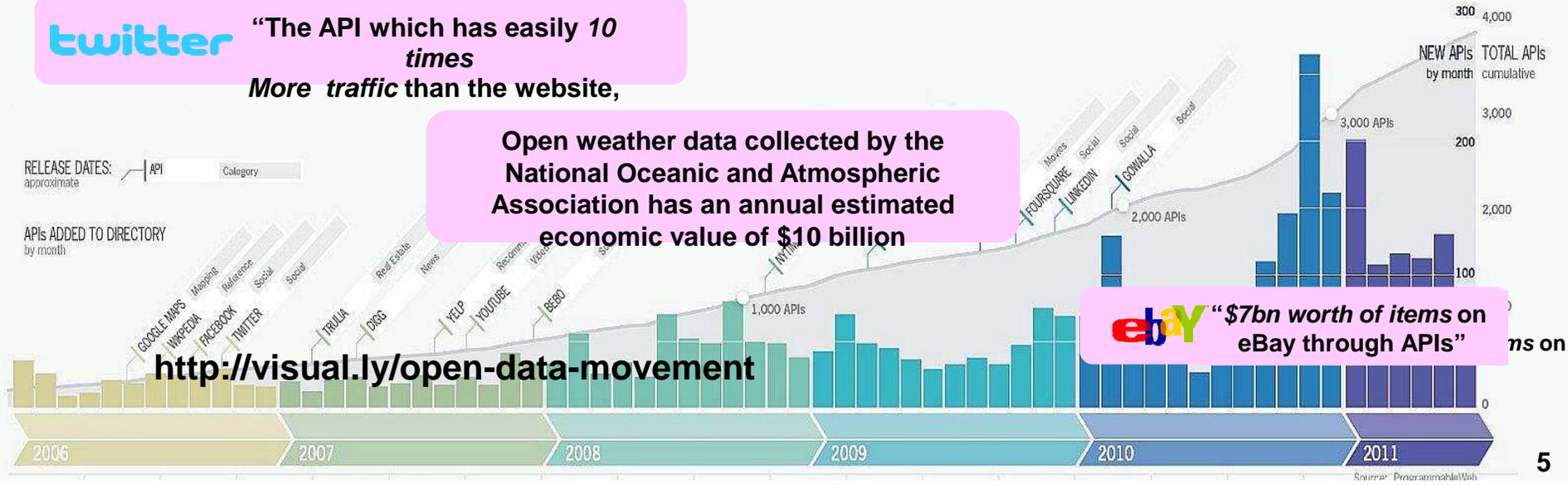


"The API which has easily 10 times More traffic than the website,"

Open weather data collected by the National Oceanic and Atmospheric Association has an annual estimated economic value of \$10 billion

 **"\$7bn worth of items on eBay through APIs"**

<http://visual.ly/open-data-movement>



Who has Open Data (partial list !)

DATA.GOV
EMPOWERING PEOPLE

Search Data Catalog Search site content

HOME ABOUT ▾ DATA METRICS ▾ OPEN GOVERNMENT ▾ BLOGS COMMUNITIES ▾

OPEN DATA SITES

Data.gov is leading the way in democratizing public sector data and driving innovation. This movement has spread throughout cities, states, and countries.

U.S. States 39	U.S. Cities and Counties 40	International Countries 43	International Regional 160
--------------------------	---------------------------------------	--------------------------------------	--------------------------------------

Download the full list of Open Data Sites in the following format:
If you would like to be part of this community and help to shape

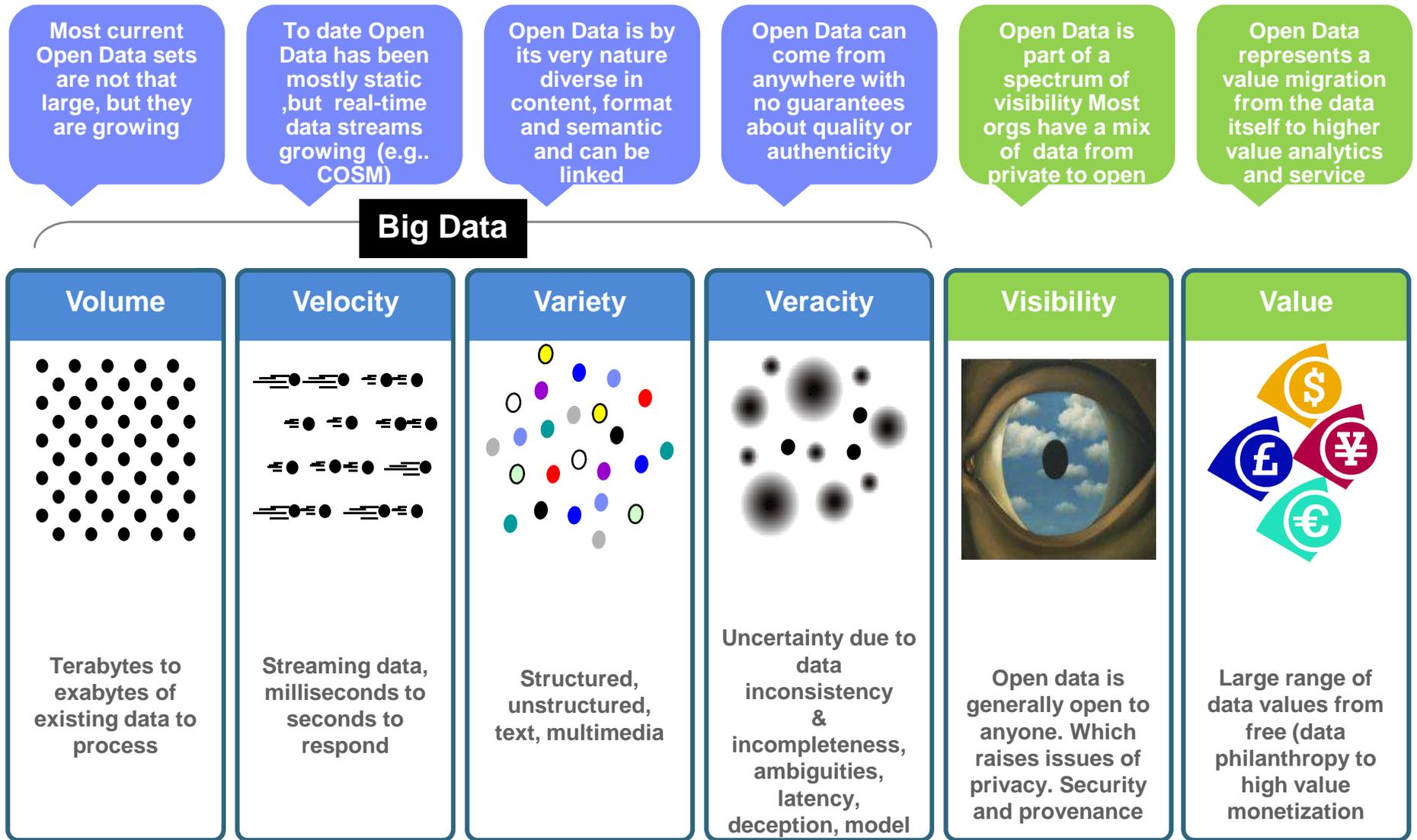


... and it's growing. Example: recent announcement by the US Federal Government

- Presumption in favor of openness for all federal data.
- Data to be made available in convenient machine readable formats...
- ... at the finest level of detail permitted by law...
- ... in a timely manner...
- ...fully described...
- ...with no restrictions on use.
- Builds on earlier announcements and also data.gov, probably already the world's largest open data collection.



Open data is 'just' data, but there are a few new considerations



Big Data + Open Data + Co-Creation = New opportunities

Common Open Data File Repositories

Portal

**420893 datasets
> 1500 Apps**

DATA.gov Releases Open Source Software

DATA AND APPS COMMUNITIES OPEN GOVERNMENT

279,829 apps and applications datasets
1,264 government apps
236 citizen-developed apps
100 mobile apps
100 agencies and subagencies
Support a dataset or app
2011 Open Government Data.gov to interactive, accessible, and social

EDUCATION ENERGY
HEALTH MANUFACTURING
OCEAN SAFETY

LEARN SEMANTIC WEB DEVELOPERS CORNER

Built On

Socrata

- Leading provider in US
- Venture-backed startup in Seattle
- Provides query, visualization, data download API
 - Over 50 customers

Platform

**8730 datasets
> 210 Apps**

DATA.GOV.UK^{Beta}
Opening up Government

Home Data Participate Apps Location Linked Data Library Lab About

BETTER OPEN DATA
SPEND DATA REPORTING TOOL
OPEN DATA CASE STUDIES
BETTER OPEN DATA
Blog by Tim Davies
OPEN DATA WHITE PAPER
Unleashing the potential
Read the UK Government's vision for the future of open data

THE NEW DATA.GOV.UK

8729 Datasets

Built On

ckan

- Leading provider in Europe
- Developed by Open Knowledge Foundation (non-profit)
- Portal software is open source
- Support and hosting available for fee

More than 35 countries now have open data repositories

Example: MetroChicago

- Aggregation of data from State of Illinois, City of Chicago and Cook County.
- Nearly 2000 interactive, explorable datasets for:
 - Public safety
 - Transportation
 - Education
 - Business services (building permits, food inspections...)
 - Healthcare...
- Chicago is an open data pioneer, in the US at least. Recently referenced by The Economist*
- Windy Grid is a new Open Source big data platform !

Search Metro Chicago Data:

Enter a term to search across data from the City of Chicago, Cook County, and State of Illinois.

MetroChicagoData.org brings interactive information and data from the City of Chicago, Cook County and the State of Illinois into one single, citizen-friendly interface. Search for keywords or browse a unified set of topics and categories.

Data Convergence

Explore Healthcare Facilities

Data for Business

Search the catalog for business-related datasets.

Data for Residents

Explore the catalogs for public safety, transportation, and education datasets.



Example: Dublin, Ireland

- Curated set of linked data for Dublin and surrounding areas, also supporting research on urban areas by a local university.
- Members collaborate to solve shared urban challenges, sharing knowledge / experience.
- Provides a trusted zone for open and proprietary data, particularly real-time operational data.
- Also supports locally created apps as well as a range of face-to-face and on-line activities.

Dublinked Datastore Apps Activities Members

Connecting a city.
Combining and sharing knowledge across disciplines, organisations, space and time. Seeking relationships that will help improve understanding of the complex interactions in a city

Welcome to Dublinked
Dublinked is a unique initiative between Dublin's Local Authorities and National University of Ireland Maynooth. This new regional data sharing initiative sees previously unreleased public operational data being made available online for others to research or reuse. All datasets are detailed on the **Dublinked Datastore**, ranging from planning, transport, environment, arts, culture and heritage. Looking for data or provide data?, simply get in

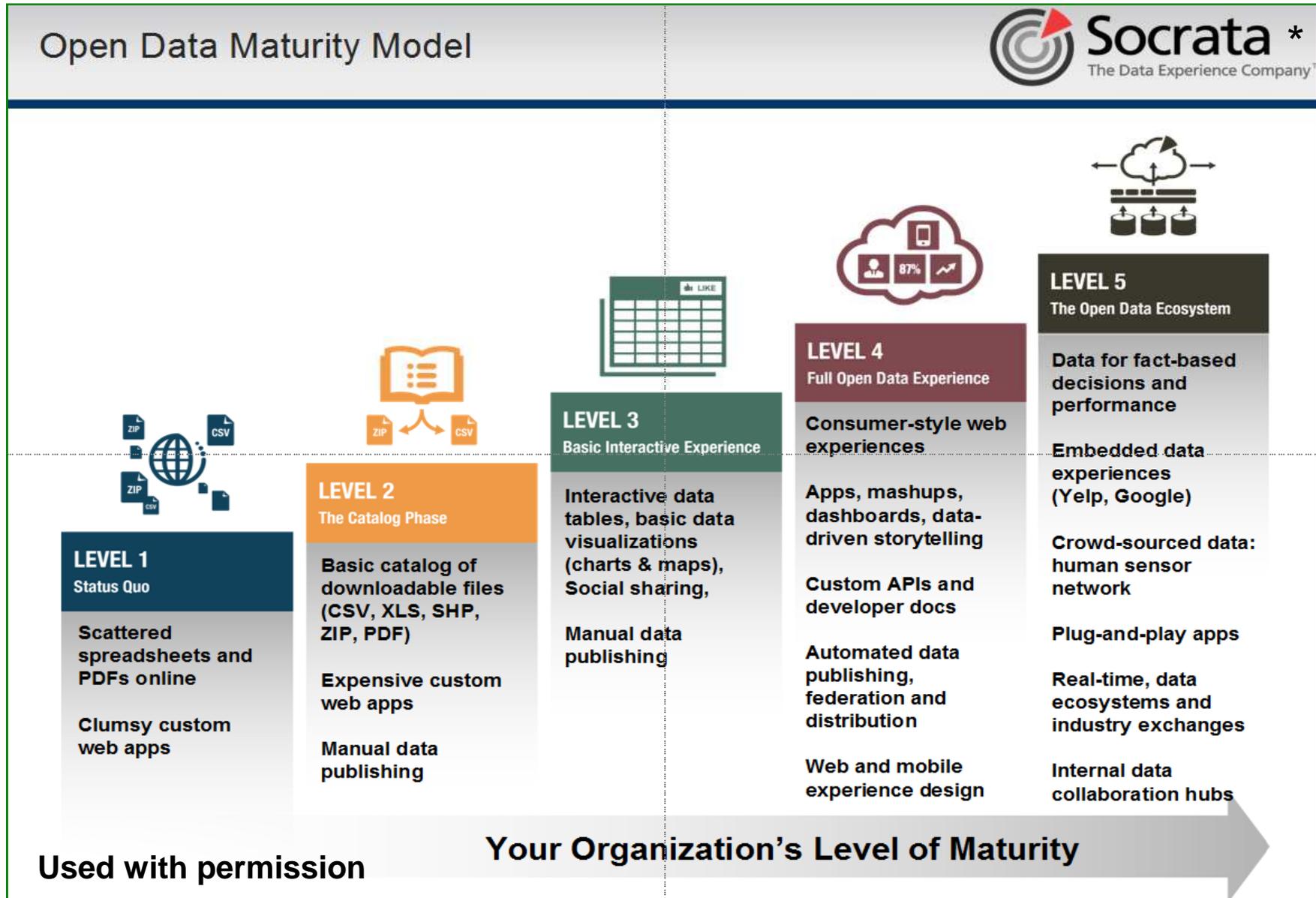
News and Blogs
'Searching in the City of Knowledge' Tutorial - SIGIR 2013
published by dublinked on Thu, 2013-05-30 14:57
As part of the 36th International ACM **SIGIR 2013 Conference** taking place in Dublin from 28th July to August 1st, IBM Research Great City Technology Centre is leading

Upcoming Events
JULYTBC: Social Enterprise DubMeet
published by dublinked on Thu, 2013-05-30 10:20
Date:
Wednesday, July 3, 2013 - 12:00
We invite interested parties to join us at our first Dublinked Innovation Network Social Enterprise DubMeet. This event will explore

TOP 10 MOST VIEWED DATASETS, FOR PAST 30 DAYS

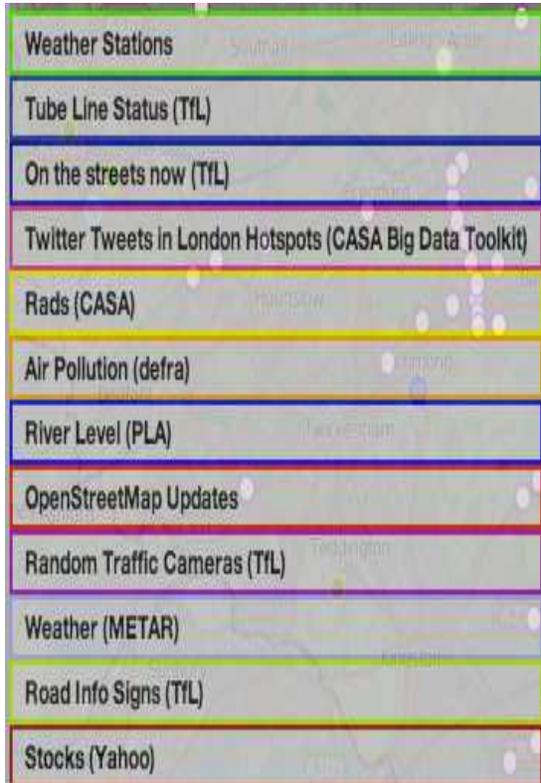
Dataset Name	Views (Approximate)
Log of Bring Centre Recycling	150
Residential parking zone information for Dublin City Area	135
Land Use Zoning maps for the 2005-2011 City Development Plan	125
Dublin City Council Administrative Areas and City Boundary	100
2011-2017 Dublin City Development Plan - Conservation Areas and Architectural...	95
Journeytimes across Dublin City, from Dublin City Council Traffic Department's TRIPS...	95
Pedestrian footfall index in Dublin city centre	85
Glass Recycling: Weights collected at Dublin City Bring Centres	75
Volume Data for Dublin City, from Dublin City Council Traffic Department's SCATS...	70
Fire Brigade and Ambulance Call Outs For 2011	70

A maturity model...



Open Data == Open Innovation *E.g. UCL's CityDashboard*

1. Source Data



✓ All data used is Open Data

2. Process

Accesses live data feeds, performing ingestion, fusion, cleansing & analytics to visualize



Uses

- Google Analytics
- Open Street Maps

✓ mix of proprietary and Open Source tools

3. Output

Visualize on Open Street Maps, using Open Layer API



✓ Open Data source

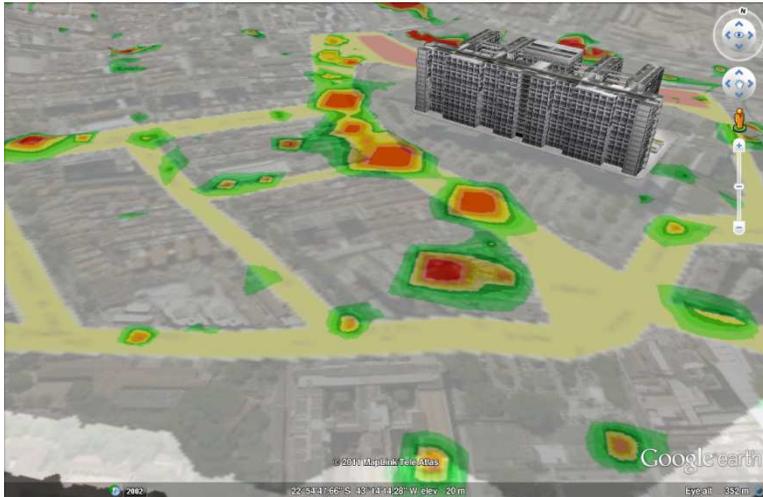
4. Package

Delivered for public use as part of a funded academic research project

citydashboard.org/london

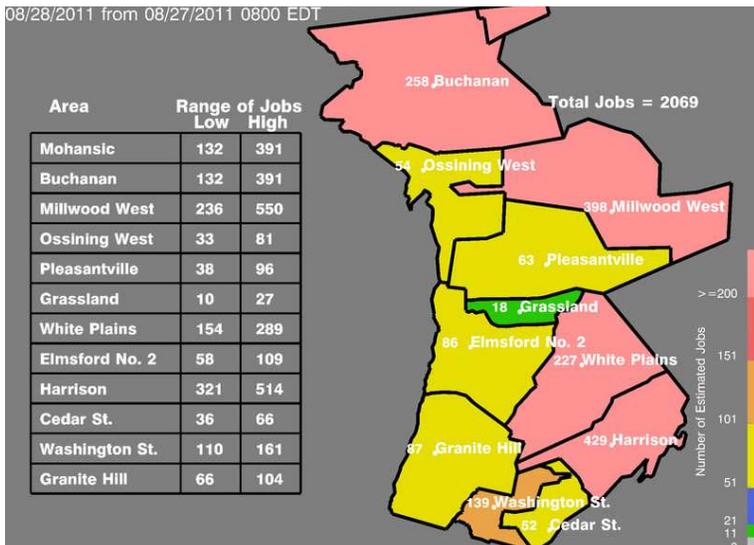
'Civic Hackers' are innovating quickly !

Example: Open data and the business of weather:



- Deep Thunder Provides customizable, accurate, high resolution weather predictions up to three days ahead.
- Models, analyses, visualises and disseminates insight from open data, organisational data, commercial data and crowdsourcing.
- Forecasts the impact of weather on businesses and systems such as energy, cities, transportation, and agriculture.
- Cities, water, energy, and transportation systems can be Smarter, safer and more efficient if the effects of weather on underlying business processes can be predicted and optimized:

08/28/2011 from 08/27/2011 0800 EDT



- **Energy:** foresee outages; predict supply from renewable resources; proactively manage load and maintenance ...
- **Cities:** foresee emergencies; proactively manage demand, water quality and maintenance ...
- **Transport:** foresee disruption to road, rail, air and marine transport; proactively manage capacity and demand; proactively manage maintenance ...
- **Agriculture:** foresee flooding and droughts; proactively manage planting, harvesting, fertilisation and supply chain activities ...
- **Others:** insurance, mining, retail, sports, tourism ...

Example: Somerville, MA - 311 Request Explorer

Somerville 311 Request Explorer

Use this tool to search, filter, and explore 311 requests that have been filed in your neighborhood. You can use the filters on the left to filter data, the geo-search box to look for requests near your home or place of business, and drill into details on each request by clicking on its title below.

Search for an address

Enter address here

Status is

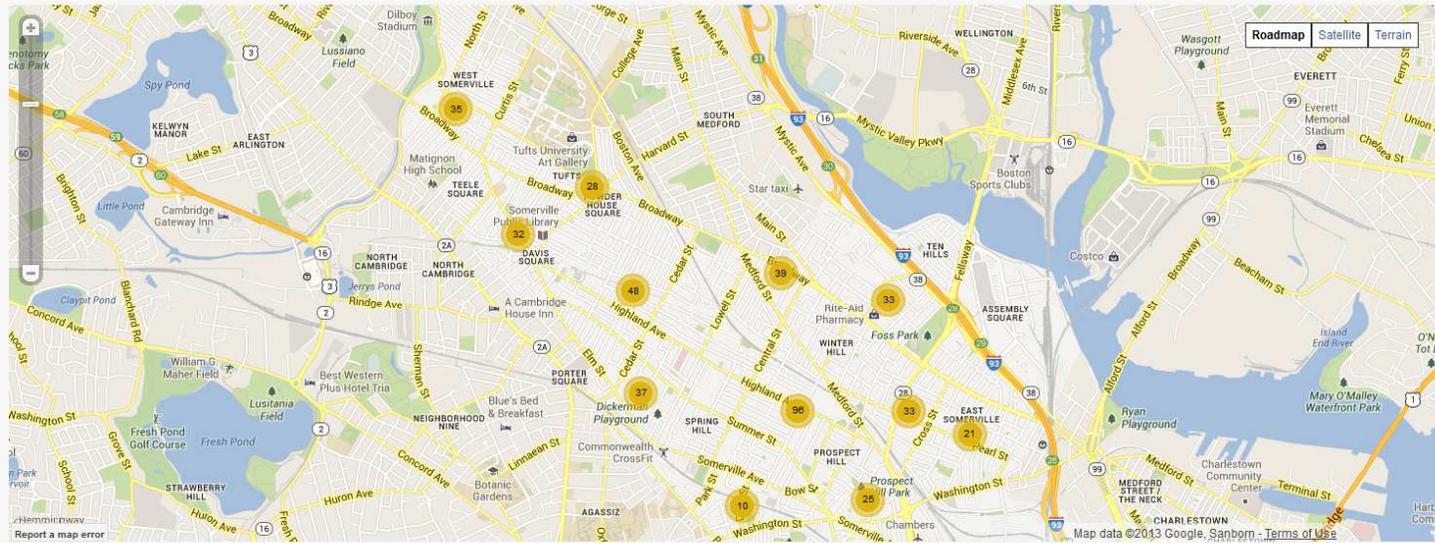
- Open
- Closed

Ticket is

Agency Name is

- Animal Control
- Comm-Connect
- Constituent Services
- DPW-B&G
- DPW-Engineering
- DPW-Highway
- DPW-Misc
- DPW-Sanitation
- DPW-Water & Sewer
- Health

500 requests match your search



#261500 **Open** **BEACON ST & SACRAMENTO ST**
 Filed: 2013-06-11T00:00:00
 Type: dataset
 Agency Name: T & P

#261368 **Open** **14 BROMFIELD RD**

Filed: 2013-06-10T00:00:00
 Type: dataset

My Take on Hackathons



The roles – We need to work together

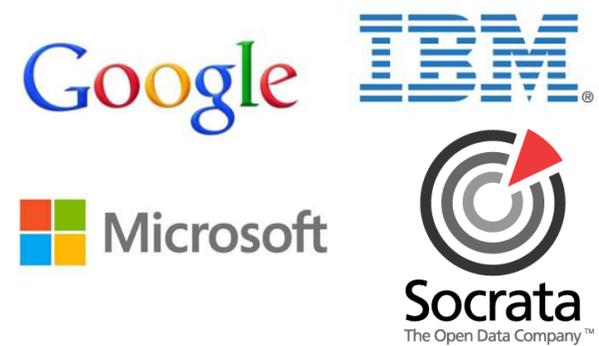
Governments



Civic Hackers



Corporations



Great opportunities for Open Data in Vermont

- Progressive and manageably sized government
- Active and engaged public
- Vibrant hacker/maker culture

- Possible areas to explore
 - Internal org transparency
 - Combined energy data (electric, gas, water)
 - Natural resource stewardship