

Big Data, Big Deal?

Alfred T. Ho
University of Kansas



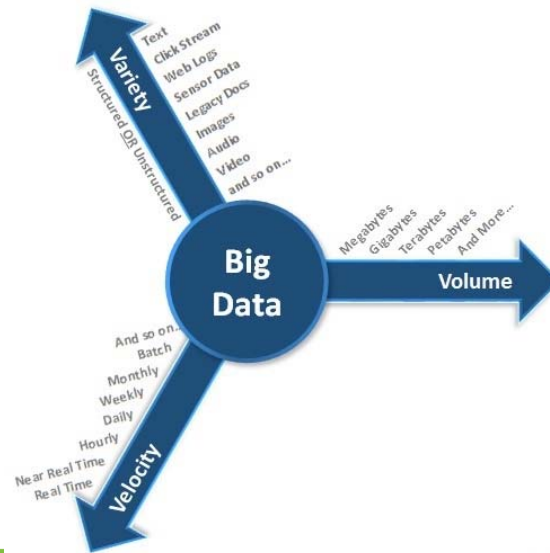
Agenda

- What is “Big Data”?
- Why is the Big Data trend relevant and important to local governments?
- Examples
- What should local managers do?

ICMA

What is “Big Data”?

Definition of “big data”
by Doug Laney (2001):
the 3 V’s



ICMA

What is “Big Data”?

- Other researchers: (Chen & Hsieh, 2014; Desouza, 2014; Ho, et al., 2015)
 - **Viscosity:** resistance to flow of data, which requires new processing to turn data into insight
 - **Variability:** changing rate of data flow
 - **Veracity:** biases, noise, abnormality
 - **Volatility:** how long data is valid for and should be stored for
 - **Vulnerability:** Greater individual risk & vulnerability

ICMA

Big Data, Big Deal?

- EMC: the world had about 800,000 petabytes of data in 2009. The amount of data are doubling every two years.
 - What's next:
 - *Exabyte* = 1,048,576 terabytes
 - *Zettabyte* = 1,073,741,824 terabytes
 - *Yottabyte* = 1,099,511,627,776 terabytes

Sources: Zdnet (2010). Size of the data universe: 1.2 zettabytes and growing fast. <http://www.zdnet.com/article/size-of-the-data-universe-1-2-zettabytes-and-growing-fast/>
Adshead, A. (2012). Data to grow more quickly says IDC's Digital Universe Study. <http://www.computerweekly.com/news/2240174381/Data-to-grow-more-quickly-says-IDC-Digital-Universe-study>

ICMA

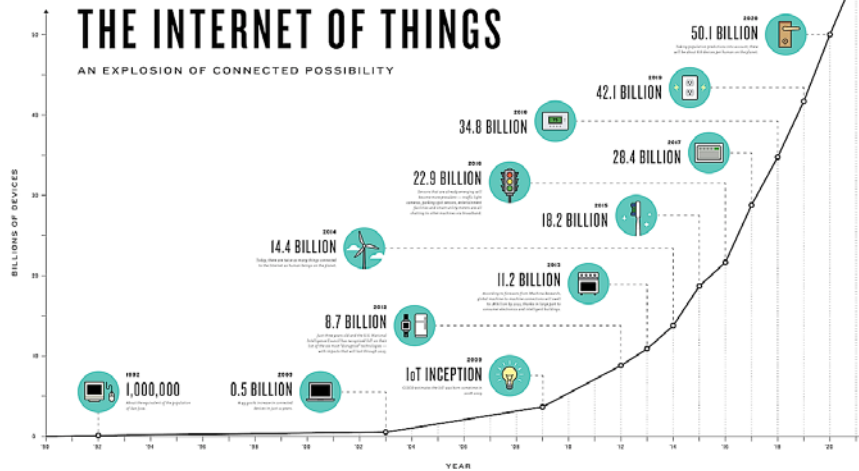
Big Data, Big Deal?

- So by 2020, every human being will each has about 5,247 GB of data



ICMA

Big Data, Big Deal?



ICMA

Relevance and Importance of Big Data to Local Governments

- First, local governments are data-rich:

- administrative records
- web info
- social media content
- communication data
- sensors
- CCTV videos
- smart meters



ICMA

Relevance and Importance of Big Data to Local Governments

- Data analytics tools can help us rethink
 - Planning
 - Customer service
 - Citizen communication
 - Resource allocation
 - Logistics management
 - Personnel management
 - Procurement
 - Performance measurement and evaluation
 - Interdepartmental collaboration
 - Intergovernmental collaboration

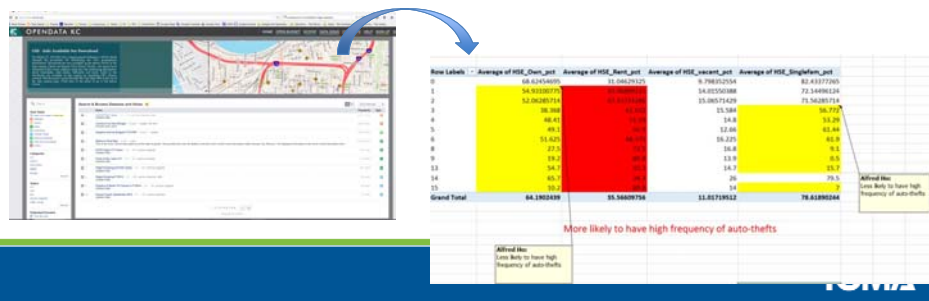
ICMA

Example

– Let's start with some “*tiny* data”

From my performance management class: Ask students to get the crime data from the “Open Data KC” platform, then focus only on auto parts theft cases (about 5,000 cases) and geocode the data spatially with census block group information, then export the data to EXCEL to see what factors are associated with the incidence of auto theft.

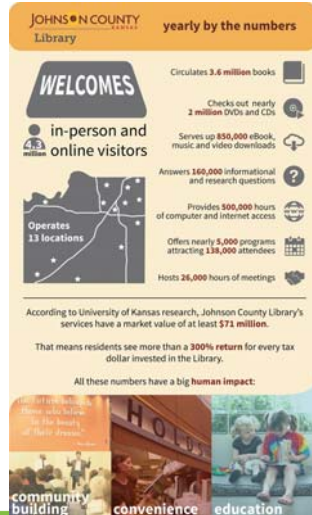
- *Guess where you are more likely to find auto parts thefts?*



ICMA

– A “*small data*” Example

(but slightly bigger than the previous example)



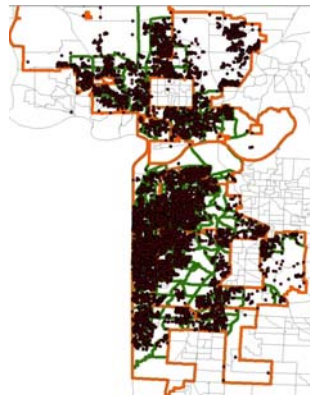
By analyzing hundred thousands of library records and using sample data from Amazon.com, we estimated the market value of **Johnson County Library** services for local residents.

We then surveyed more than 9,000 library patrons about their willingness to pay and various preferences, and integrated the data with the Library’s administrative records, such as programming and attendance records, to evaluate library patron preferences by demographic groups and by library branches.

The information is now being used to rethink the Library’s strategic direction and programming needs.

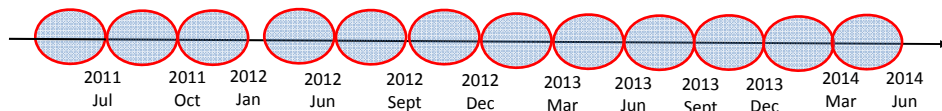
ICMA

– A “*Semi-Big data*” Example



Using more than 13,000 KCMO citizen survey responses in 2011-2014, 90,000+ citizen service requests every year, about 60,000 police cases every year, and census block group information, a research project conducted by the KU team used various statistical models to analyze the factors that were associated with citizen satisfaction with KCMO services and quality of life issues

→ Rethink strategic priorities and program focus



ICMA

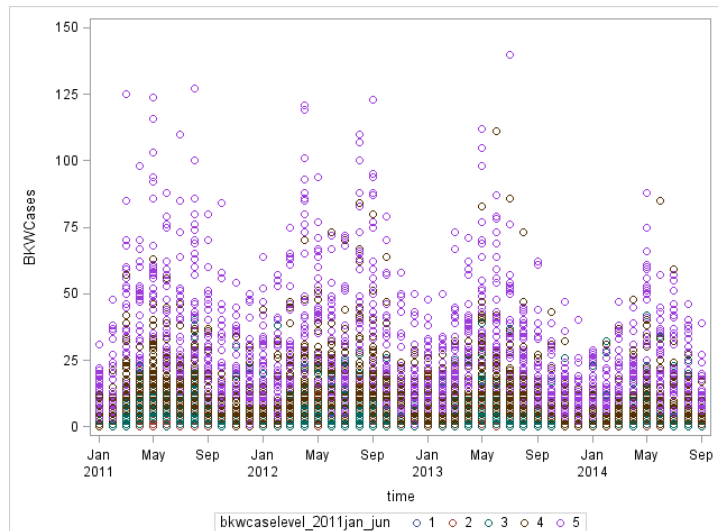
– A “*Semi-Big* data” Example – KCMO

- By clustering hundreds of thousands of 311 complaints, we found that these cases tend to cluster: graffiti, illegal dumping, vacant properties, property nuisance

ICMA

– A “*Semi-Big* data” Example – KCMO

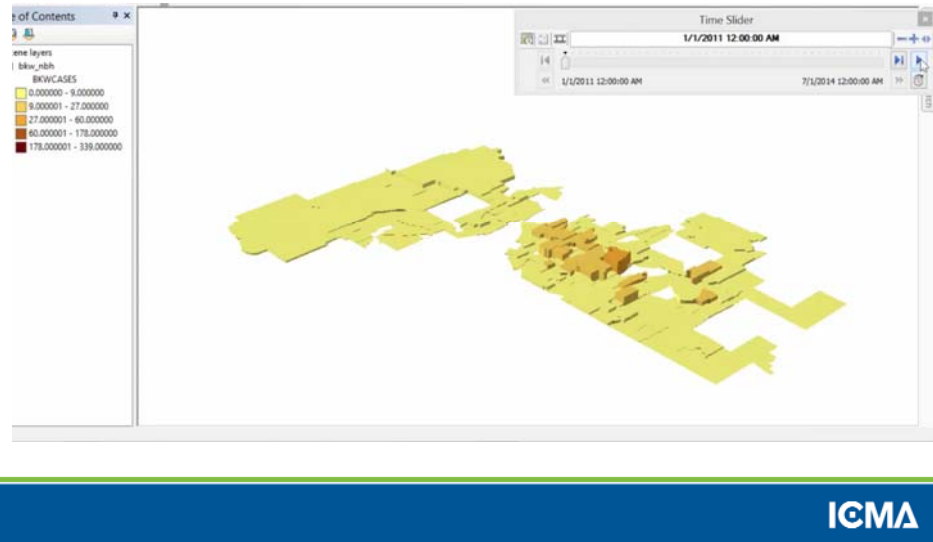
But big data analytics allows us to probe more deeply:



ICMA

– A “*Semi-Big* data” Example – KCMO

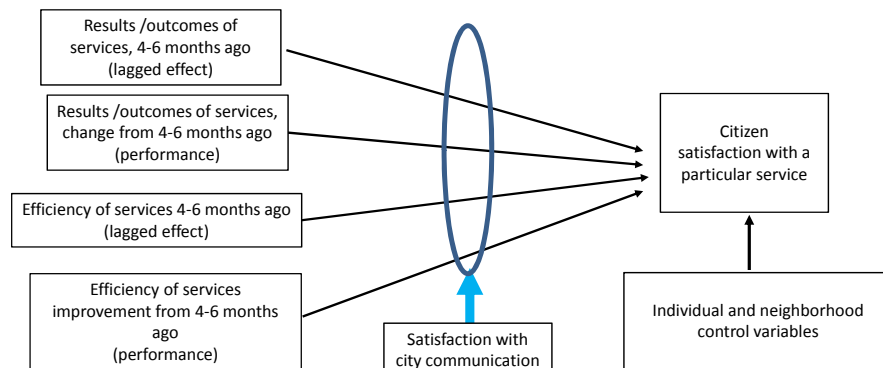
- You may also show the data in another way ...



ICMA

– A “*Semi-Big* data” Example – KCMO

- What is most valuable about data analytics is the additional insights about the association of factors and the possible impacts of policy actions:



ICMA

– A “Semi-Big data” Example – KCMO

A surprising finding:

Dependent variable: Satisfaction with Police Protection

	Evaluating at the means or medians	Communication effectiveness = "Dissatisfied"	Communication effectiveness = "Highly dissatisfied"	Crime rate = mean + 1*SD	Crime rate = mean + 2*SD
Highly dissatisfied	2.36%	4.20% (+1.84%)	7.37% (+5.01%)	2.41% (+0.04%)	2.45% (+0.09%)
Dissatisfied	7.09%	11.72% (+4.62%)	18.18% (+11.09%)	7.21% (+0.11%)	7.32% (+0.23%)
Neutral/Don't know	33.86%	42.10% (+8.30%)	45.97% (+12.11%)	34.15% (+0.29%)	34.44% (+0.58%)
Satisfied	46.95%	36.31% (-10.64%)	25.30% (-21.65%)	46.66% (-0.29%)	46.37% (-0.58%)
Highly satisfied	9.73%	5.61% (-4.12%)	3.18% (-6.56%)	9.58% (-0.16%)	9.42% (-0.31%)

Source: Cho & Ho (2015)

The more dissatisfied residents are with city communication, the more likely that they are also dissatisfied with police protection, after controlling for many demographic & neighborhood factors.

Surprisingly, the substantive impact of crime rates on perception of police protection is relatively small

ICMA

– A “Big Data” Example

- The “Street Bump” app by the City of Boston

The screenshot shows the City of Boston website. The main content area features a section for the "Street Bump" app. The app is described as a mobile app that helps identify potholes by recording "bump" data. A "Get the App" button is visible, along with a search bar and a "News & Press Releases" section.

ICMA

A Predictive Analytics Example - the City of Chicago



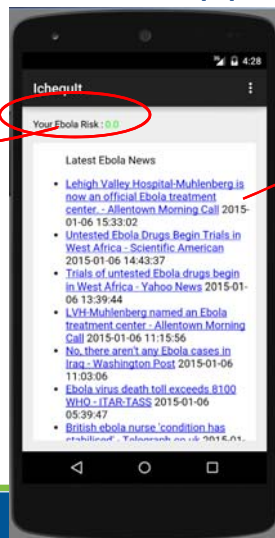
If you had to list one creature that is least welcome by any city dweller, a rat would be a safe bet. *Rattus norvegicus*, otherwise known as the brown rat, Norway rat, or common street rat, inhabits every continent except Antarctica, and particularly thrives wherever dense concentrations of people exist. Wherever they may roam, rats commonly carry and spread diseases, pose a threat to food supplies, and—because of their constantly growing teeth—cause damage to wires, pipes, and other infrastructure. In short: rats must be stopped (or at the least, contained).

In partnership with the Event and Pattern Detection Laboratory (EPD Lab) at Carnegie Mellon University, Chicago's Department of Innovation and Technology (DoIT) is taking on a predictive approach to the "war on rats." By using data in innovative ways to help keep rat populations down, Chicago is putting to use a new strategy that can not only enhance rodent control initiatives,

ICMA

Another Predictive Analytics Example - A KU-invented App: ichequit

Individualized risk assessment based on CDC data and social media information



News and information Feeding based on an individual's location



Good privacy protection due to phone-based risk calculation

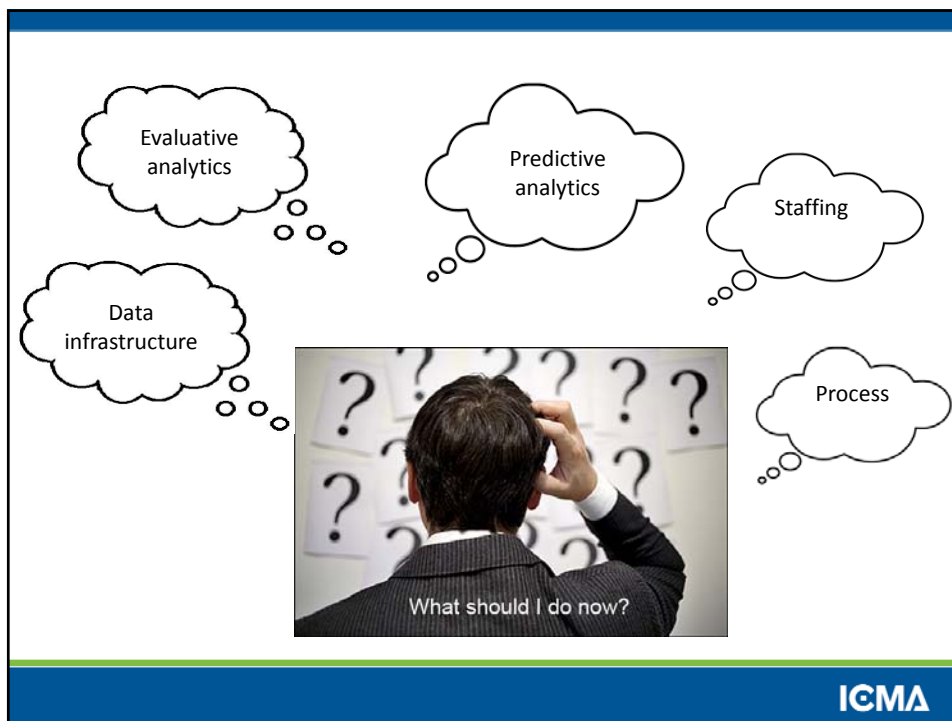


ICMA

Many other examples ...

- Many cities are using Big Data to analyze crime patterns and pursue “predictive policing”
- New York City is using Big Data to analyze homeless problems so that they can take proactive steps to help needy families
- Many public works departments and local utilities are using Big Data to find cost savings and vulnerabilities
- Many auditors and revenue departments are using Big Data to find errors, fraud, and illegal transactions

ICMA



ICMA

What Should Local Managers Do?

- Focus on the Needs of Citizens and Clients
 - Big data can't just focus on technologies and equipment
 - It should focus on the needs of the public and clients
 - What service quality do they expect?
 - How can data analytics used to serve their needs?
 - What are the implications on citizens and service clients
 - What are their concerns?



ICMA

What Should Local Managers Do?

- Develop a Hierarchy of Data Strategies

	Open data	Semi-open data	Internal data
Who can decide? What processes are used?			
How privacy and security measures are used?			
Who are the intended users?			
Purpose of collecting /using the data			
Time cycle of policies			

ICMA

What Should Local Managers Do?

- Need to Spend Resources on Analysis

- Lots of data \neq useful intelligence

- Policy relevance?
- Appropriate data scope and depth?
- Appropriate tools?
- Concerns about data quality?



- Correlation \neq Causal relationship

- We still need careful research design, data strategies, and appropriate analytical tools to evaluate policy outcomes and estimate causal impacts

ICMA

What Should Local Managers Do?

- Develop Network Strategies

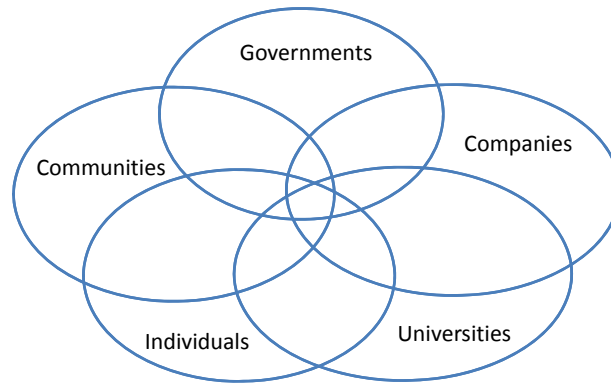
- Can't rely on a single department or even a single unit of government to do this well
- Need a network of data partners and collaboration strategies



ICMA

What Should Local Managers Do?

- Think about Governance, not just Government



ICMA

What Should Local Managers Do?

- Develop Ethical Guidelines for Data Policies and Strategies

- Not just think about what is legal, but what is ethical
- Is a practice appropriate and acceptable by the community?



Perhaps a role for the ICMA in the future?

ICMA

What Should Local Managers Do?

- Become more Security Conscious
 - The more data we have, the bigger the security concerns become
 - This can be a significant capacity challenge for local governments
 - Pressure to pursue contracting-out?



ICMA

Conclusion: Big Data, Big Deal?

- Yes, because big data is already an institutionalized phenomenon
 - Big data is changing how we live our daily lives
 - Big data is changing politics
 - Big data is changing citizen expectations



→ *Inevitably, big data will change the world of local management*



ICMA

Conclusion: Big Data, Big Deal?

- We need to embrace these *CHANGES!*
 - Citizen or client focus
 - Hierarchy of data strategies
 - Analysis
 - Network strategies
 - Governance thinking
 - Ethical practices
 - Security consciousness

ICMA

Questions/Comments?

Contact information:

Dr. Alfred T. Ho
Associate Professor
School of Public Affairs & Administration
University of Kansas
Phone: 785 864-3416
Email: alfredho@ku.edu



