Big Data – better business benefits

Paul Edwards, HouseMark

2 December 2014
What I’ll cover……..

• Explain what big data is
• Uses for Big Data and the potential for social housing
• What Big Data means for HouseMark
• Some practical data protection considerations
• What I won’t be doing:
  ▪ Giving detailed legal advice and guidance - there are other better qualified people doing this today…….
What is Big Data?
What is Big Data?

• We hear a lot about big data and its capacity to transform our daily lives. But what do we mean by Big Data and what does it mean in practical terms for social housing providers and their customers?
What is Big Data?

There are lots of definitions of big data - & some of them are incredibly complex and technical.

‘a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications’ - Wikipedia

‘data sets with sizes beyond the ability of commonly used software tools to capture, curate, manage, and process the data within a tolerable elapsed time’

*Like drinking from a firehose*
The Four Vs of Big Data

- **Volume (Amount)**
  - The quantity of data to be captured continues to grow exponentially.

- **Variety (Type & Sources)**
  - Data comes in many formats, from diverse sources.

- **Velocity (Speed)**
  - Bits and bytes have to be processed at high speed.

- **Veracity (Quality & Trust)**
  - Data needs to be converted into meaningful insights.
Big Data

In simple terms:

- **Big Data** is a popular term used to describe the exponential growth and availability of data, both structured and unstructured.
- So there are ever increasing amounts of data out there and in lots of different forms.
- There’s a view that **Big Data** may be as important to business – and society – as the Internet has become. **Why?**
How might Big Data be used?
The hopeful vision is that organisations will be able to take data from any source, harness relevant data and analyse it to find answers that enable, for example:

- Cost reductions
- Time reductions
- Smarter business decision making
- Better service
- New product development and optimised offerings
Big Data in Action

• Exploiting the "four virtual Internets" individually and collectively:
  ▪ People
  ▪ Things
  ▪ Data
  ▪ Ideas

• They are emerging to enable broader collaboration and knowledge and also as an invaluable source of big data to fuel advances in business capabilities
The Internet of People

- **Social networks**
  - opportunities to create better and more instant communication and feedback with customers through social media

- **Crowd sourcing**
  - Street Bump is a crowd-sourcing project that helps residents improve their neighbourhood streets. Volunteers use the Street Bump mobile app to collect road condition data while they drive. The data provides governments with real-time information to fix problems and plan long term investments.
The Internet of Things

- Exponential increase in intelligent objects connected to the internet
- By 2020 estimates suggest 50 billion devices connected to the internet, as smart sensors become more present
- Massive potential here for social landlords
  - Care for the elderly
  - Energy efficiency
  - Health & Safety
  - Asset management
The Internet of Data

- Increase in data storage and processing capability creates huge potential for data mining & analysis
- Opportunity to bring together key data sources
  - Key intelligence to identify need and plan services
  - More intelligent management of people and assets
  - Tailoring services around individual needs
  - Real time data to support performance management and business planning
  - More sophisticated forward predictions – e.g. asset lifecycles and costing
The Internet of Ideas

- The Galaxy Zoo project used volunteers to classify over a million galaxies
  - An example of citizen science – building capacity to make sense of the mass of available data

- Open Notebook Science makes the full record of scientific research available online
  - This enables other scientists to obtain detailed descriptions of procedures, raw and analysed data to either compare with their own work or to build on

- Scope for social housing to share research which underpins the development of evidence based practice?
Big Data for all

Big Data is not just for big organisations

• Cloud partnering advantages
  ▪ cost
  ▪ scalable
  ▪ new capabilities
  ▪ avoid data silos
  ▪ security and compliance aspects are solvable

• All businesses need other data
  ▪ Open data
  ▪ Third party data
  ▪ Private sector companies swap customer data, e.g. credit reference agencies, supermarkets and suppliers
Big Data can also be small data

- Small timescales: services based on real-time sandbox analytics not just massive scale processing
- Small segments: segment-of-one personalisation is what a consumer values the most
- Big Data is also really small, sub-second or ultra personalised information
- … and every level in between
- … it’s complicated understanding every customers’ life
Analytical Strategy: Structuring Unstructured Data

- Unstructured data just means **patterns we haven’t found yet**.
- Geospatial location data, voice, video, streaming data, behavioural data.

- Example: **changing frequencies of buying behaviours** – identify/influence actionable customer segments
- which people are more sensitive to influencing at certain reoccurring times?
- the reverse: remove them from promotions if they are least sensitive at certain times
Analytical Strategy: Structuring Unstructured Data

• Example: Segmentation based on **serial basket journey analysis**
  
  ▪ how did a VIP customer become a VIP customer?
  ▪ who can be nudged into becoming a VIP?
  ▪ segmentation based on serial behaviour & predictive models

Potential big spender?
... or problem tenant?
... or expensive property?

2 new projects

Images: Wikimedia Commons
Analytical Strategy: personalisation

Big Data can also enable:

- **Personalisation of production**
  - automation: advice, suggestions, support, queries, tips
  - more precise use of resources
    - more granular campaigns
    - Tailoring offers to local circumstances

- **Personalisation of delivery**
  - channel, technology, tone, timing
  - by person
  - by minute/day/time period

- **Personalisation requires customer data because**
  - Customers need different things
  - Customer needs are dynamic
Analytical Strategy: Visualisation

• Perhaps the most important thing is when it comes to Big Data is that visualisation encourages interaction with the data.

• Studies show that managers using visualised data are more than twice as likely as their peers to interact extensively with it - 33% vs. 15%.

• Exploring data visually lets the data’s story unfold vividly in a way the brain can grasp in a flash. “A light bulb goes off,” says Wells Fargo Vice President of Strategic Planning Dana Zuber, he says, “You just don’t get that with a spreadsheet.”
Personalised Insurance - Telematics

- Black box technology
- Tracks acceleration, braking, cornering, journey length and time
- Dynamic and individual calculation of premiums

Big data benefits

- More flexible products
- Targeting services
- Incentivises behaviour
- Competitive advantage
What does Big Data mean to HouseMark in practical terms?
Big Data opportunities

• Big Data provides opportunities to:
  ▪ Adopt new ways of using data in and between organisations
  ▪ Connect the fundamental business objectives back to the raw data using data science analytical techniques
  ▪ Increase acceptance and expectation of strategy and operations based on analysis and evidence not just gut feel
Big Data analytics

- Big Data analytics involves the triangulation of data sources to:
  - discover patterns
  - to create business intelligence
  - to inform decision making
- Big Data analytics involves:
  - structured data – HouseMark benchmarking
  - unstructured data – emails, website logs and comments
  - open data – DCLG, HCA, LAs, Home Office, ONS
  - commercial data – Ordnance Survey, RICS, Dataspring
  - machine data – GPS vehicle data, energy meters, appliance
What does Big Data mean to HouseMark

- It’s about being aware of and utilising:
  - The latest available data sources
  - The latest tools and techniques for mass data
    - Capture
    - Storage
    - Interpretation and analysis
  - Conveying insights to a wide audience
A Housing Data Ecosystem

Data coverage and sourcing

- Income & cost
- PI & satisfaction
- Assets & stock
- Treasury & funding
- Tenants & residents
- Census
- Market rents
- IMD
- Housing need
- House prices

Member data

External data

Standard
Custom
Bespoke
Toolkits & profiling
Data sharing

Products, outputs and reporting
Opportunities for Big Data in social housing

- Challenging fraud
- Improving health and safety in buildings
- Addressing fuel poverty
- Enhancing asset management
- Incentivising rent payments
- Tackling social exclusion
Now
• Massive dataset capturing landlords’ costs, performance and satisfaction.

Next
• More ‘real-time’ collection of existing data.
  ▪ People and property database pilot
• Link to open data and other sources.
  ▪ Open Data Zone
Data capture

Now

- Strong reputation for bringing consistency to data collected from over 400 landlords through clear definitions and rigorous validation.

Next

- More collection direct from system source – ‘automated data collection’.
- Advances in collection methods open up opportunities of greater granularity in data collected – right down to property and tenant level if necessary.
- Capturing data from new sensor technology in social housing – the ‘Internet of Things’
Now

• In-house data analysts within HouseMark’s Data Services team.

Next Academic partnerships

• Partnerships with university data scientists and statisticians, starting with the Advanced Data Analytics Centre at Nottingham University.
Now

• Traditional paper reports and web-based presentations of data, including dashboards and scorecards with a VFM focus.

Next

Cutting-edge tools

• Extending GIS capacity in conjunction with funding from the Nottingham Geospatial Institute.

• More use of Microsoft business intelligence and analysis reporting tools on our interactive website.
Some practical data protection considerations
Data protection considerations

• Not all big data analytics involve personal data
  ▪ E.g. climate & weather data

• Personal data
  ▪ Is it necessary? Can it be anonymised?

• Privacy impact assessment
  ▪ Personal data - to identify general trends or make decisions that affect individuals

• Repurposing data
  ▪ Is the new purpose compatible with the original? Is consent required?
Data protection considerations

• Fair processing
  ▪ Transparency re collection and purpose

• Information security
  ▪ Lost data
  ▪ Stolen data: Climategate, Edward Snowden
  ▪ Privacy, confidentiality, etc.
    • Jigsaw effect - if you know the sex, date and year of birth, and the city of someone in the US, then 53% of the US population can be uniquely identified
Data protection considerations

• Data minimisation
  - Big data analytics not an excuse for stockpiling data – can you articulate/justify long term uses?

• Transparency
  - Be transparent & open
Some tips & learning points……

• Get proper advice
• Be clear what is and what is not personal and sensitive data
• Be clear what you want to do with the data – clear purpose…. 
• Ensure it is securely stored

• It’s more complicated than you might think!!
Summing up
Key Considerations

• Don’t forget about your own data
  ▪ Make better use of your own data
  ▪ Supplement with openly available public data

• The technology is not the big problem, it’s skills —
  ▪ big data technologies are widely available
  ▪ the big problem is having the right skills to exploit it
  ▪ Being better informed is not much use unless you can make good use of the better information
  ▪ Is there a sector solution to this?
Big Data in Action

- It’s not about the data – it’s about the insight it gives you to underpin your decision making, increase your operational efficiency and improve outcomes

- Big data can help you to:
  - Design services and customer access channels
  - Balance service demand and service quality
  - Make best use of available resources
  - Inform investment and manage risk

- Private sector will
  - Exploit big data to maximise profit

- Public sector challenge
  - To use big data to maximise social value from available resources
Thanks for listening

Paul Edwards
paul.edwards@housemark.co.uk
07841 861974