The promises and pitfalls of big data

Dr Dianne Heriot
Parliamentary Librarian
Parliament of Australia
Overview

• Big Data v Big Information
• Promises – what does it offer Parliamentary Libraries?
• Pitfalls – what can go wrong?
• Navigating the issues
Big Data v Big Information

Big Data:

“Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization”
Big Data v Big Information

• **Large data sets** (n=everything): mobile phone locational data, number of downloads

• ‘Messy data’: multiple, sometimes unstructured data sets of variable quality

• **Probabilistic**: correlations between variables – not specific details nor causation
Big Data v Big Information

Big Data
- Lots of detail about similar things in no particular order

Big Information
- Large amounts of information about individual topics
Promises – what does it offer Parliamentary Libraries?

Information

- Purchasing
- Analysing and advising
- Legislation

Understanding our clients
- Downloads and page views
- Text analysis
Pitfalls – what can go wrong?

*Not to scale
Pitfalls – what can go wrong?

- Data quality
- Data formats
- Procurement risk
Navigating the issues

- Project Management
- Skills development
End