The Data Cure or the Data Curse? Exploring the Yin and Yang of Big Data

Kerrie Mengersen
Professor of Statistics
Faculty of Science and Engineering
Institute for Future Environments
Australian Research Council Centre of Excellence
Mathematical & Statistical Frontiers:
Big Data, Big Models, New Insights

7 year horizon
6 Universities
7 Partner Organisations
50 Professorial Investigators, >50 Researchers&students
The Big Data Charge

Figure 1
Data is growing at a 40 percent compound annual rate, reaching nearly 45 ZB by 2020

Data in zettabytes (ZB)

Source: Oracle, 2012

Exabyte: one billion billion = one quintillion bytes
The Big Data Capability

Exabyte: one billion billion = one quintillion bytes
The Big Data Capacity

The Number of Internet Connected Devices Surpassed the Global Population in 2010

Internet Connected Devices / People (bn) vs. Devices per Person

- Connected devices
- World population
- Connected devices per person

Source: Company data, Morgan Stanley Research
The Big Data Cost

Figure 1 - Source: Wikibon 2014

The Big Data Challenge

Big Data Investments by Industry

Has your organization already invested in technology specifically designed to address the big data challenge?

- Education: 28% Yes, 17% No plans at this time, 12% No, but plan within 2 years, 4% No, and 39% Don't know
- Retail: 12% Yes, 18% No plans at this time, 29% No, but plan within 2 years, 21% No, and 29% Don't know
- Transportation: 4% Yes, 21% No plans at this time, 11% No, but plan within 2 years, 15% No, and 36% Don't know
- Healthcare: 9% Yes, 15% No plans at this time, 15% No, but plan within 2 years, 25% No, and 36% Don't know
- Communications, Media, Services: 10% Yes, 15% No plans at this time, 21% No, but plan within 2 years, 17% No, and 21% Don't know
- Insurance: 14% Yes, 28% No plans at this time, 11% No, but plan within 2 years, 18% No, and 31% Don't know
- Energy, Utilities: 5% Yes, 28% No plans at this time, 28% No, but plan within 2 years, 18% No, and 22% Don't know
- Banking: 18% Yes, 18% No plans at this time, 8% No, but plan within 2 years, 18% No, and 23% Don't know
- Manufacturing: 13% Yes, 37% No plans at this time, 19% No, but plan within 2 years, 12% No, and 23% Don't know
- Government: 16% Yes, 35% No plans at this time, 19% No, but plan within 2 years, 18% No, and 23% Don't know

Source: Gartner (July 2012)
The Big Data Cure
A common challenge is ensuring patient data security
The Big Data Curse

Are Agencies Taking The Steps To Prepare?

One in three say their agency has successfully launched at least one Big Data initiative

Only 34% have invested in IT systems/solutions to optimize data processing

29% have trained IT professionals to manage and analyze Big Data

29% have educated senior management on Big Data issues

Fewer than one in five say their agency is very prepared to work with Big Data

Big Data Diagnosis

Despite challenges, 9 out of 10 respondents expect M2M technologies and data to have positive impacts across the healthcare industry
The Yin and the Yang of Big Data

“The data are big!”
The Yin and the Yang of Big Data

“Garbage in, Garbage out”
The Yin and the Yang of Big Data

“Now that we’ve got it, what do we do with it?”
A statistician’s data cue
Why Big Data?
Because sometimes the solutions are not where we are looking