



# MarkLogic Briefing

Providing Customer Value Through  
Data-Led Services

The Data Revolution and its Consequences

David Worlock

November 2011



# What is Big Data?



- Data collections that are difficult to handle and search in conventional terms because of their size.
- Is that all? What does it mean?
- The exhaust fumes of the networked society?  
The dandruff of bureaucracy?

# The Lexis Risk Management Case Study



**Size?** Experian, Choicepoint and the US Government

**Target?** The US insurance industry

**Result?** Users contribute their data, scheme becomes industry intensive, and begins to exhibit compliance/due diligence characteristics

# Astra Zeneca Case Study

Prescription information/medical-claims data + clinical evidence = ?

- Evidence-based focus
- Partnership with customers (WellPoint/HealthCore Feb 2011)
- Which diagnostics make sense and which don't?
- Which medical devices and which drugs?
- What tends to errors and/or high readmission rates?
- What are the local as well as the systemic issues?

# Sentiment Analysis



- Xaxis (WPP - a "data and marketing services company")
- Initial target - the buying decisions of 500 million people in Europe, the USA and Australasia - including all UK citizens
- “Xaxis streamlines and improves advertisers ability to directly target specific audiences, at scale and at lower cost than any other audience -buying solution”
- Also note the data alliance of SAP and Google

# Evidential Research



- The Nature/Signalling Gateway experiment
- Replacing the article with the evidence
- Searching science in 3D - abstract, article, evidence
- Multiple media evidence - and the importance of taxonomy, ontology and semantic discovery

# B2B plays everywhere



- Environmental Risk - from EDR to Landmark at DMGT
- Oil Exploration - Elsevier, IHS and Wood Mackenzie collaborate at GeoFacets
- Engineering - Globalspec file 10m design briefs
- Aviation Ascend - All the parts that are fit to fly

# Financial Services



- Machine readability
- News and the Celerity environment
- Cross file search necessity - DataExplorers
- Thomson Reuters - Open Calais/ClearForest



# All of the tools are at your disposal

- Hadoop/HortonWorks/Cloudera/HCCP (Lexis)
- Most of the lessons were learnt in Google, Yahoo!, Facebook
- How do you stitch it all together? MarkLogic
- How do you participate - by taking an additional revenue contribution from your data content?
- Or by becoming a hub in a virtual data connection designed to create new answers in a field where you have authority, customers, distribution, or a new vision of the unanswered questions...

# Thanks for listening!



Contact David Worlock at  
[david@davidworlock.com](mailto:david@davidworlock.com)

Tel: 07836 361873

On Twitter at [dworlock](#)

And come to [www.davidworlock.com](http://www.davidworlock.com)