

Training on Big data for Asia and the Pacific

8 December 2017

Miriam Hodge, PhD



Definition

- The 3 V's: volume, variety, and velocity
 - Organic Data
 - Data not collected for statistical purposes
-
- It has been called the new oil. “A new commodity spawns a lucrative, **fast-growing industry**, prompting anti-trust regulators to step in to restrain those who control its flow.” – The Economist

Big Data UN Global Working Group

The Statistical Commission agreed at its 45th session (2014) to create the Global Working Group (GWG) on Big Data for Official Statistics



4TH INTERNATIONAL CONFERENCE ON
BIG DATA
for Official Statistics

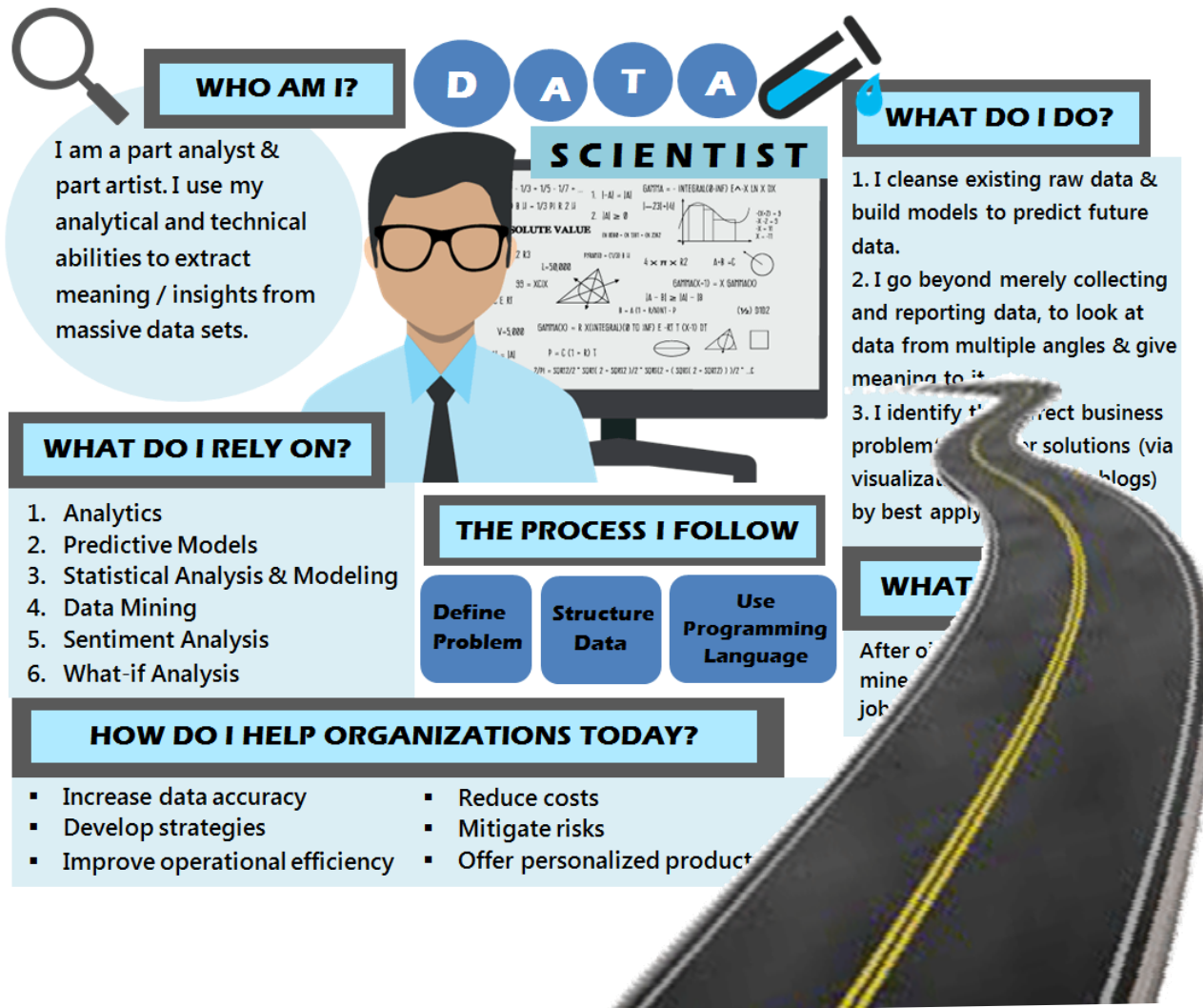


8-10 NOVEMBER 2017
BOGOTA, COLOMBIA

More than just data

- Techniques for analyzing data, such as data mining, machine learning and natural language processing
- Big data technologies, like business intelligence, cloud computing and databases
- Visualization, such as charts, graphs and other displays of the data

DATA SCIENTIST ROADMAP (BPS)



Data Science Campus (ONS - UK)

The goal of the Campus is to build a new generation of tools and technologies to exploit the growth and availability of innovative data sources and to provide rich informed measurement and analyses on the economy, the global environment and wider society.

Mission statement:

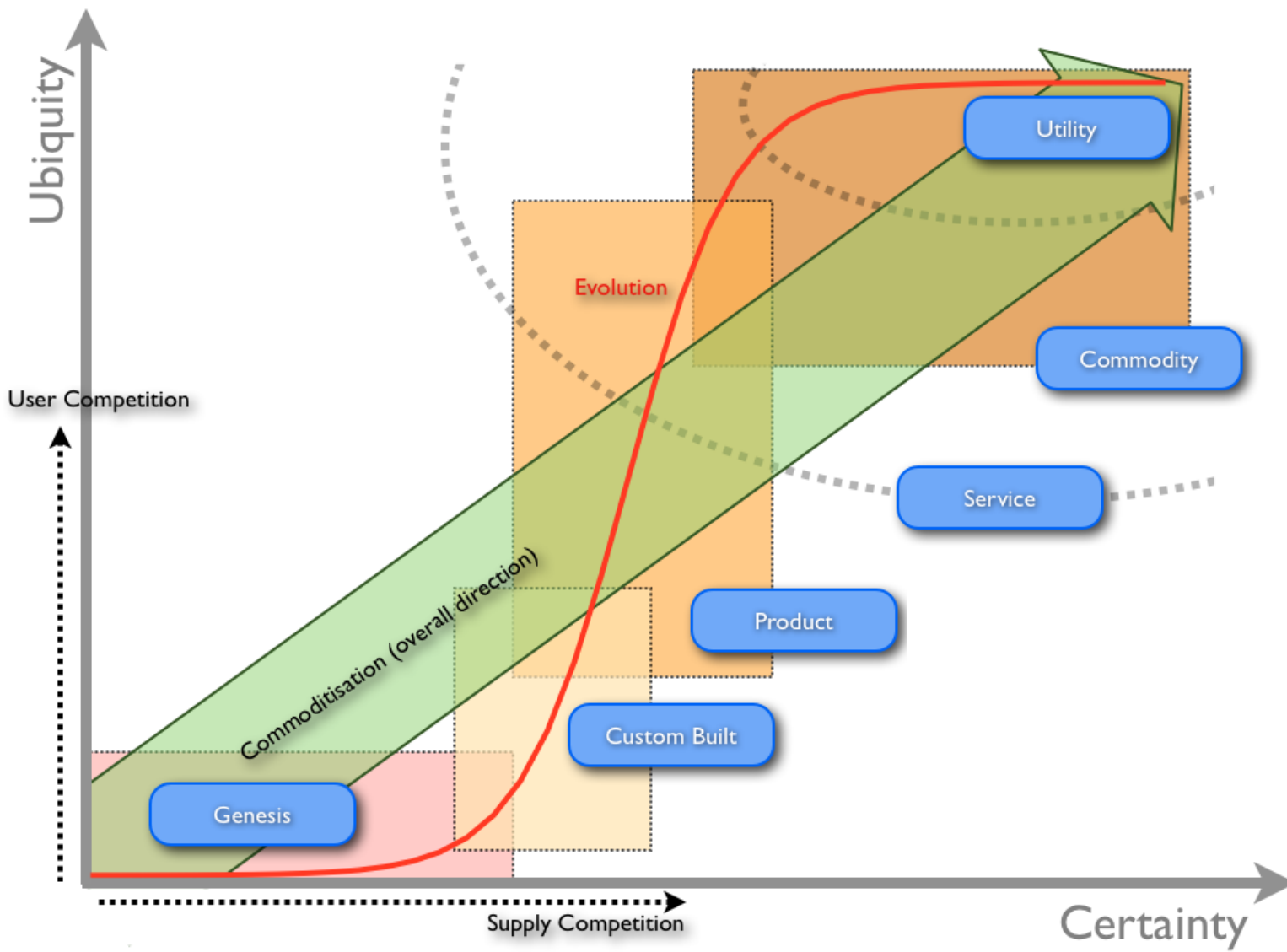
‘Data Science for the public good’

The goal is to adopt a new mindset within ONS.
Only survey for benchmarks and filling gaps.



<https://datasciencecampus.ons.gov.uk/>

The process of evolution repeats ...



SIAP

- Satellite Imagery and Geo-Spatial Data
 - Geospatial data module
- Social Media Data
 - Web scraping
 - Collecting Twitter data
- Technologies
 - R Studio
 - Python
 - QGIS

Creating technical training material

- Technical skills training is time intensive
- Not just lectures
- Developing reference material with step-by-step instructions or training videos
- Training facilities with computers, networks, and training platforms that can handle big data.
- Regular updating of material to keep up with rapid changes

Conclusion

- Training Data Scientists and bringing their skills into National Statistical Offices is an essential part of using Big Data for Official Statistics

References

- <https://unstats.un.org/bigdata/>
- <https://www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource>
- Manyika, James; Chui, Michael; Bughin, Jaques; Brown, Brad; Dobbs, Richard; Roxburgh, Charles; Byers, Angela Hung (May 2011). ["Big Data: The next frontier for innovation, competition, and productivity"](#). McKinsey Global Institute. Retrieved January 16, 2016.