



It's about Place not Space

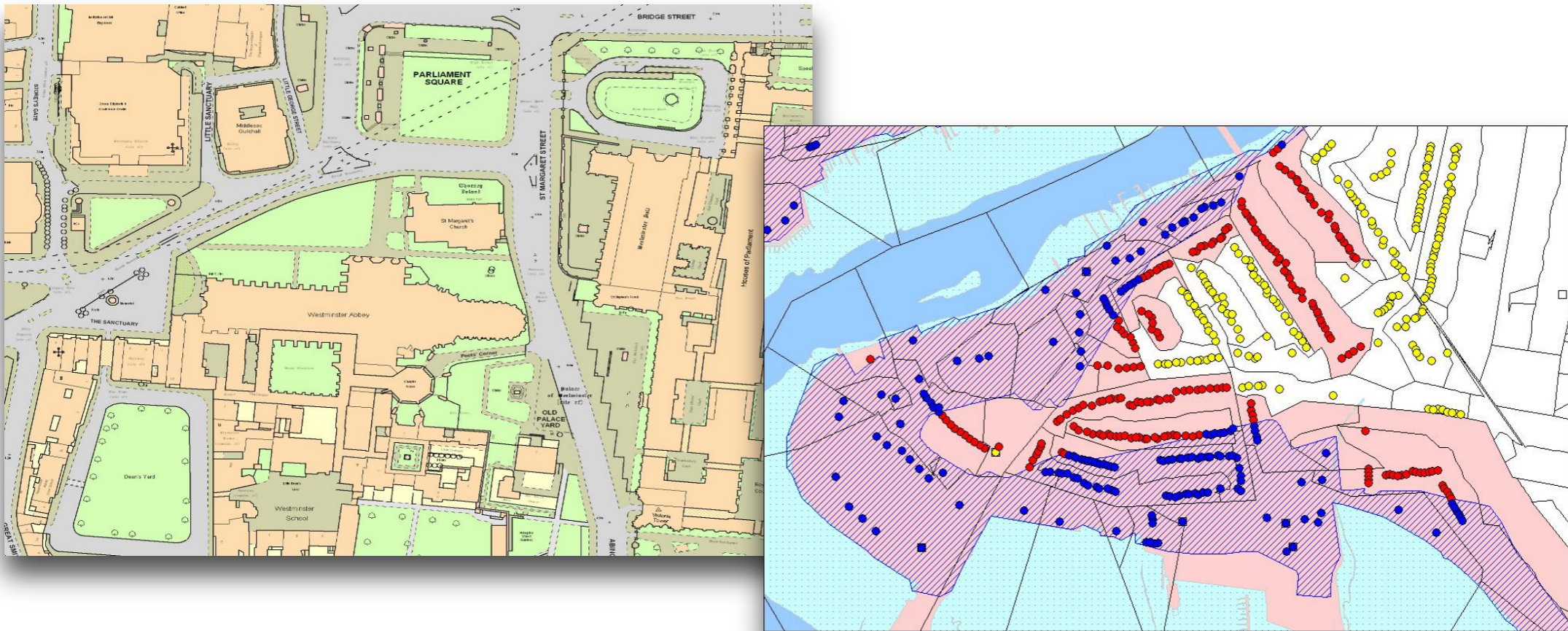
IBM Big Data, Integration and Governance Conference

Matthew White
Business Development Consultant

6 November 2012

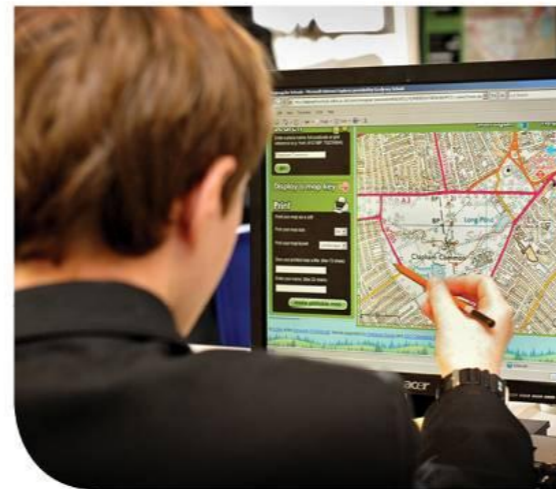
Geospatial

- Everything happens somewhere
- Geospatial is a key driver of decision-making for public and private sectors
- Over 80% of structured and unstructured data has a geospatial element



Ordnance Survey Great Britain

- Ordnance Survey is 220 years old
- Civilian organisation since 1983; 1150 staff
- Independent Government Department and Executive Agency reporting directly to a Government Minister
- Trading Fund since April 1999
- Annual Report for 2010/11: Revenue of £129.4m, profit before exceptional items of £24.1m, dividend £6.3m
- Headquarters in Southampton with 28 field offices around Great Britain

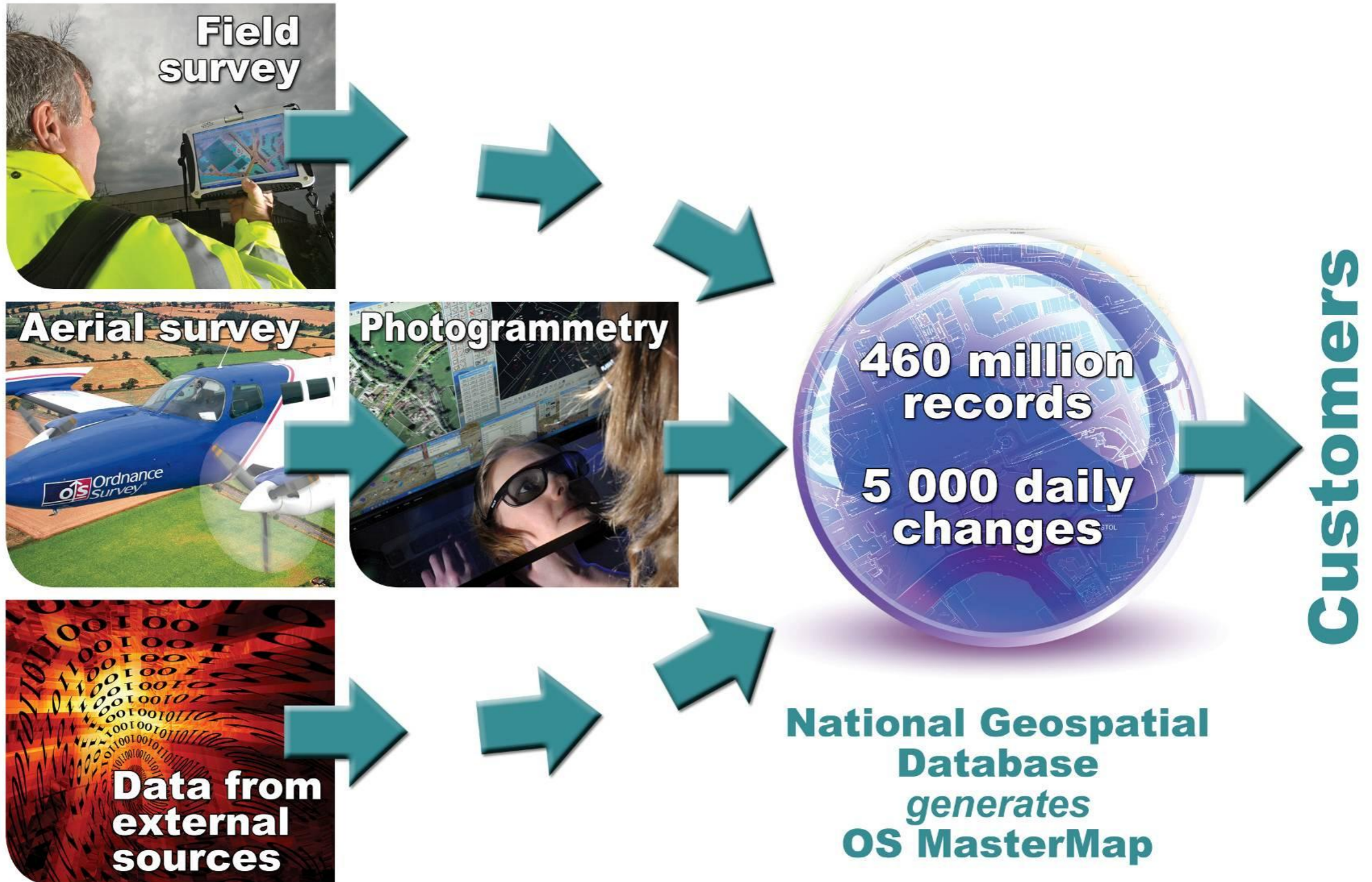


Ordnance Survey today

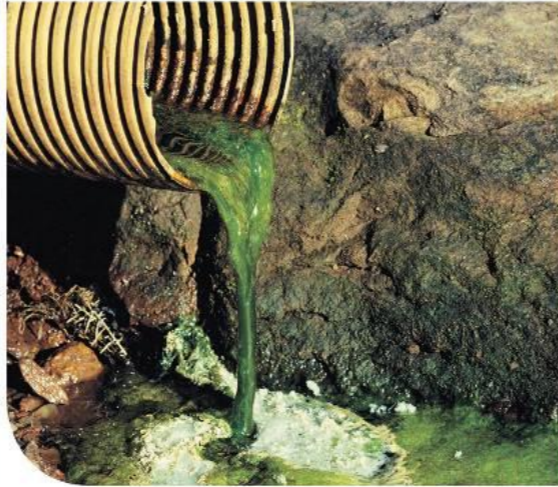
- Creates and maintains the 'master location database' of Great Britain
- Maintains a database of 460 million features
- Approximately 5,000 changes made daily
- In 2010/11, 99.9% of real world features were represented in the database within six months of completion on the ground
- From the database, Ordnance Survey produces a range of geospatial data products for business, leisure, educational and administrative use



Updating the Ordnance Survey database



Wide Range of Customers and Markets





BROADWAY

TOPHILL STREET

THE SANCTUARY

St James's Park Station

Queen Anne's Chambers

St Anne's Chapel (site of)

Dept for International Development

Department for Business, Innovation & Skills

Dean's Yard

Blackstone Tower remains of

Westminster School

New Scotland Yard

VICTORIA STREET

Abbey Orchard Street

Sanctuary Buildings

Park House

Clark House

St Edward's Chapel

PO

Bank

TO 6s

Playground

Archite Centre

Club

Community Centre

47 to 50 Westminster Mansions

Griffons (PH)

1 to 12

1 to 32 C

1 to 24

Alec Wilzard House

St Mathew's School

Edward Atrop Court

St Anne's Church

1 to 25

1 to 14

St Ermin's Hotel

1 to 10

1 to 25

1 to 24

St Anne's Church

St Mathew's Church

1 to 25

1 to 14

Griffons (PH)

1 to 12

1 to 25

1 to 24

St Anne's Church

St Mathew's Church

1 to 25

1 to 14

Griffons (PH)

1 to 12

1 to 25

1 to 24

St Anne's Church

St Mathew's Church

1 to 25

1 to 14

Griffons (PH)

1 to 12

1 to 25

1 to 24

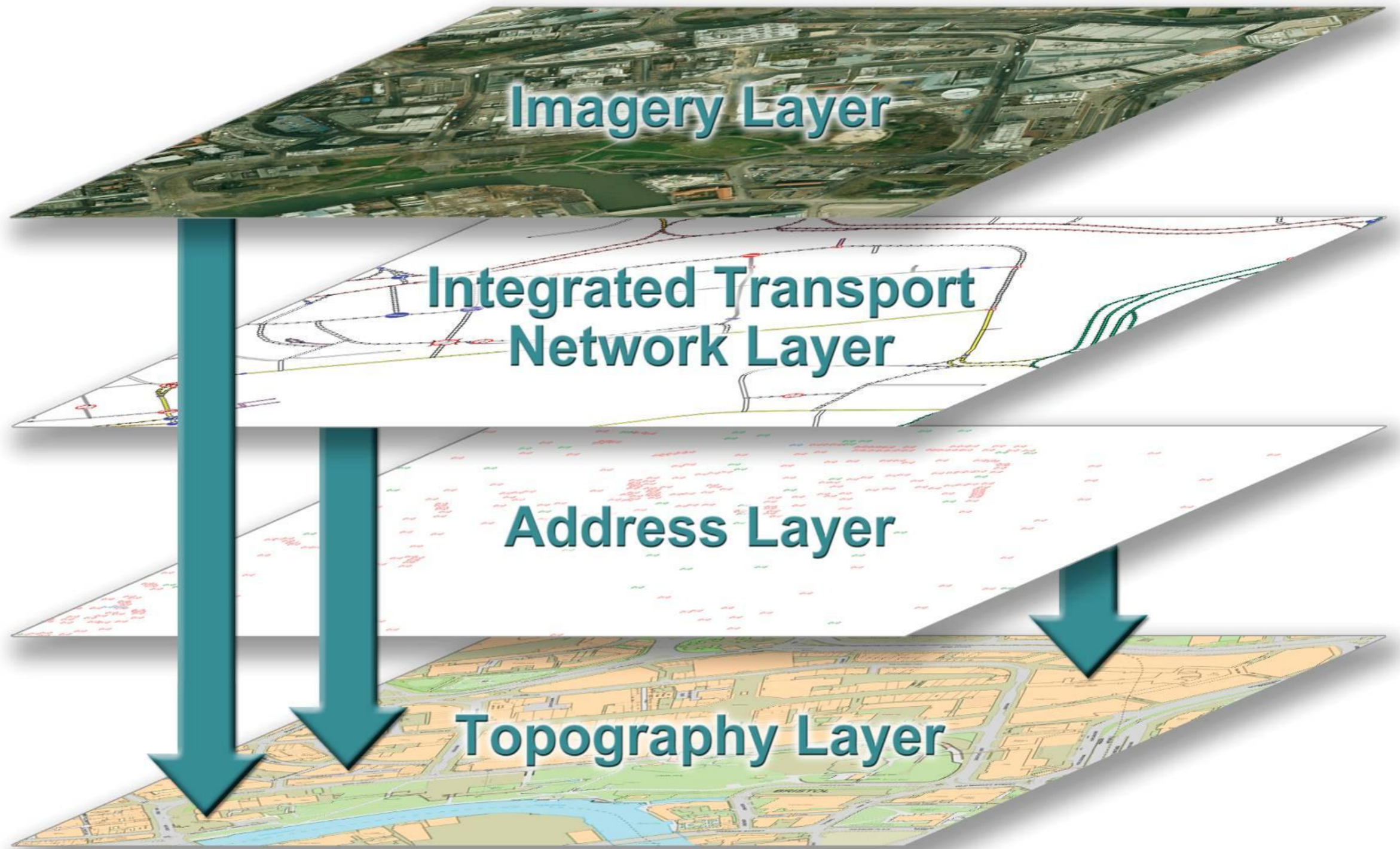
St Anne's Church

St Mathew's Church

1 to 25

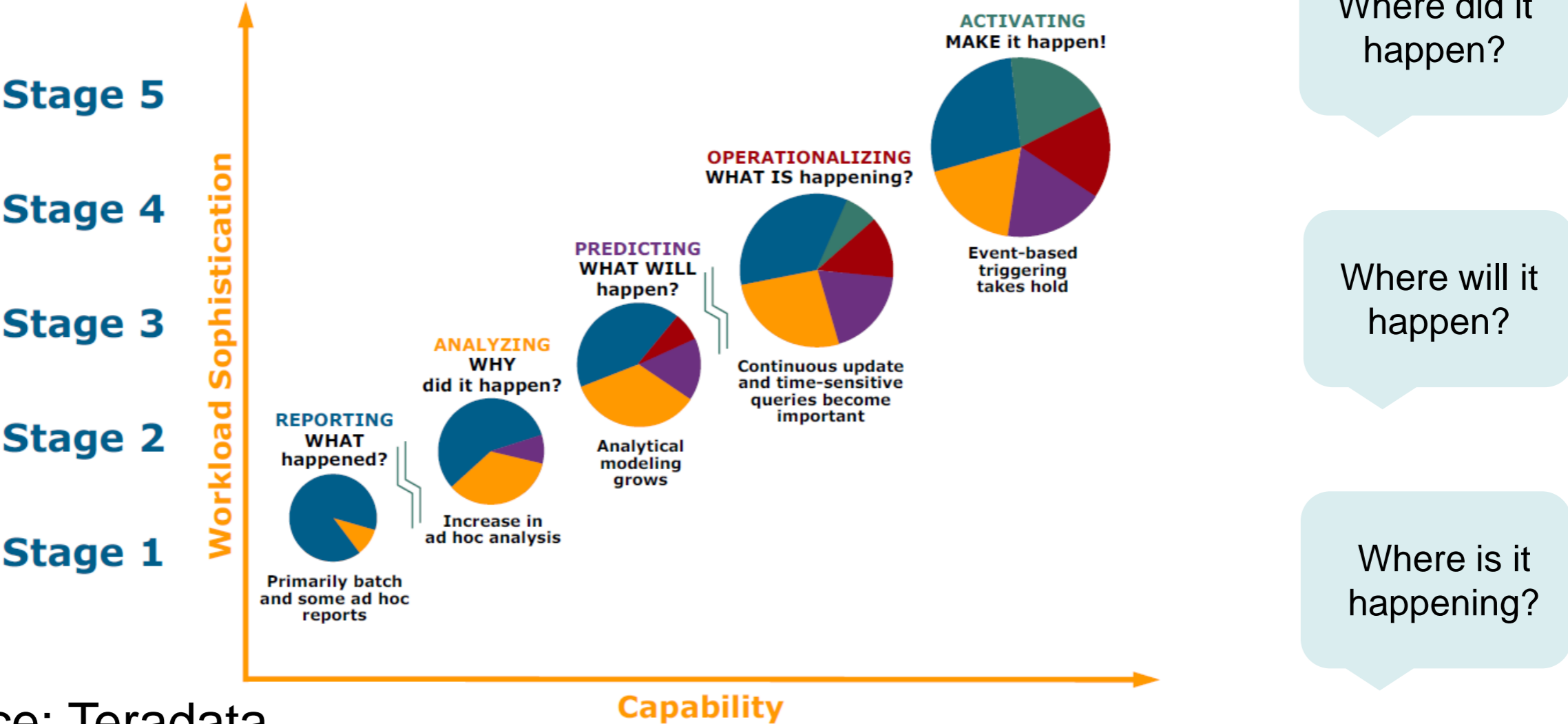
1 to 14

OS MasterMap



Analytics plus geospatial data is changing the way we get insights

- Geospatial analytics gives you the ability to ask “where” questions of business data



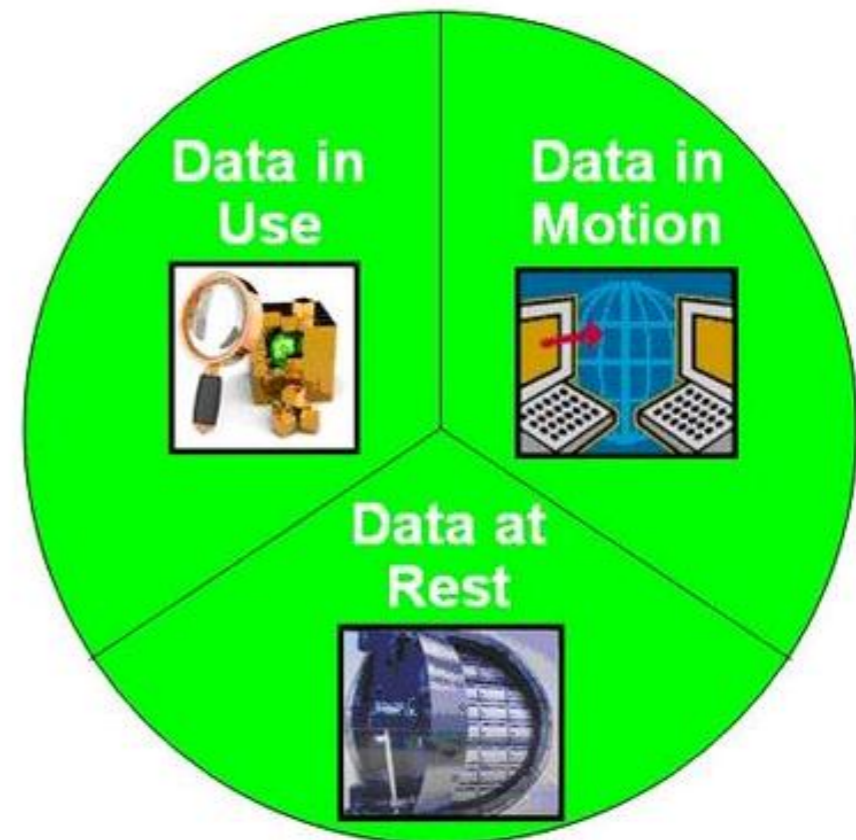
Source: Teradata

Analytics plus geospatial data is changing the way we get insights

- Where are my customers?
- Where are my competitors?
- How far will customers travel to a branch or store?
- Which of my competitor's customers can I draw to a branch or store?
- Which customers live close to a branch or store?
- Where can I increase profitability?
- How can I mitigate financial risk from flooding?

The big data environment

- Volume
 - Data at rest
- Velocity
 - Data in motion
- Variety
 - Data in many forms
- Geospatial data is all of these



Ordnance Survey and Netezza



Using Netezza for high performance geospatial analytics

**Stress Testing
our Data**

Data Queries

New Insights

**Storytelling with
Location Data**

Netezza and geospatial analytics

- In-database geospatial analytic functions
- Native understanding of geospatial data
- High performance out of the box
- Scales to terabytes of data
- No indexes or aggregates to manage
- Open, standards-based interface and data model



Analyse all data in a single appliance

Stress Testing our Data



Stress Testing our Data – Volume of Data

Internal Analytics	Status	Current	Netezza
The Address Layer 2 queries listed in AI_M2_INTERFACE_DATA_SQL.sql (which test for any invalid combinations of attributes in the dataset).	Improve	~3 hours	17 mins
Topology check of OS MasterMap Topography Layer Topo Lines and Topo Areas.	New	N/A	3 mins 3 secs
Address Layer 2 features that are matched to Topography Layer buildings (TOID reference) but do not sit spatially within them.	New	N/A	25 mins
Features within each of the OS MasterMap layers which reference a feature that no longer exists in the product (e.g. Address Layer 2 "Reference to Cartographic Text TOID" to Topography Layer "TOID").	New	N/A	6x queries at ~1 min each
Count of OS MasterMap Topography Layer Topo Lines below 5cm (by length split at 1mm intervals).	New	N/A	37 secs
Number of OS MasterMap Topography Layer Topo Lines with a Descriptive Term of "Inferred Property Closing Links" which don't have a Topo Area with a Descriptive Term of "Multi Surface" on either side of them.	New	N/A	36 hrs

? ? ? ?

Data Queries

? ? ? ? ?



Data Queries – Volume of Data

External Analytics	Status	Current	Netezza
What length of the GB coastline is made up of "beach"?	Improve	~2 weeks	4 mins 26 secs
What is the remotest point (point farthest from a metalled road) in the country?	Improve	~1 day	27 mins 22 secs
What is the farthest point from the coastline?	Improve	~1 day	9 mins 12 secs
What is the number of each Address Layer 2 base function falling within each of the GOR boundaries?	Improve	~2 days	1 min 9 secs
Locate potential high rise blocks of flats (using OS MasterMap Address Layer 2 and Topography Layer).	Improve	~3 days	49 secs

Data Queries – Volume of Data



We analysed 41 million records in 19 hours. We could not run this query in the past.

$f(x) = x^2$
 Find the derivative

$\frac{y_1 - y_0}{x_1 - x_0} = \frac{g(x+h) - g(x)}{(x+h) - x} = \frac{g(x+h) - g(x)}{h}$

$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$

$f(x) = \lim_{h \rightarrow 0} \frac{(x+h)^2 - x^2}{h}$

$= \lim_{h \rightarrow 0} \frac{x^2 + 2xh + h^2 - x^2}{h}$

$= \lim_{h \rightarrow 0} \frac{2xh + h^2}{h}$

$= \lim_{h \rightarrow 0} (2x + h)$

$= 2x$

$f'(x) = 2\sqrt{x}$

$f(x) = \lim_{h \rightarrow 0} f(x+h)$

New Insights

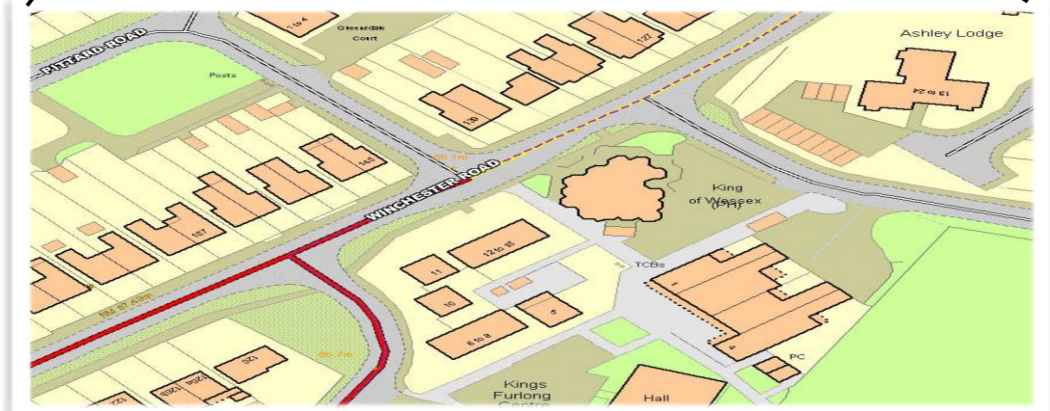
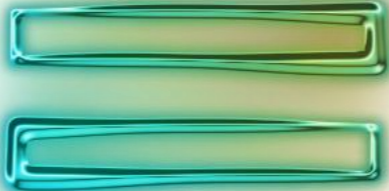
New Insights – Volume and Variety of Data

External Analytics	Status	Current	Netezza
Counts the numbers of postcodes [Code-Point]; addresses [Address Layer 2]; and addressable building polygons [Topography Layer] that fall with the Environment Agency flood polygons.	New	N/A	2 hrs 28 mins 24 secs



Storytelling with Location Data

Storytelling with Location Data – Volume and Variety of Data



Conclusions

**Stress Testing
our Data**

Data Queries

New Insights

**Storytelling with
Location Data**

Hal Varian, Google, Chief Economist

“The ability to take data – to be able to understand it, to process it, to extract value from it, to visualise it, to communicate it – that’s going to be a hugely important skill in the next decades”

Data is indeed the new Intel Inside...