Department for
Communities and
Local Government

CSaP ANNUAL CONFERENCE

Getting insights from data: from analytics to artificial intelligence - a departmental perspective
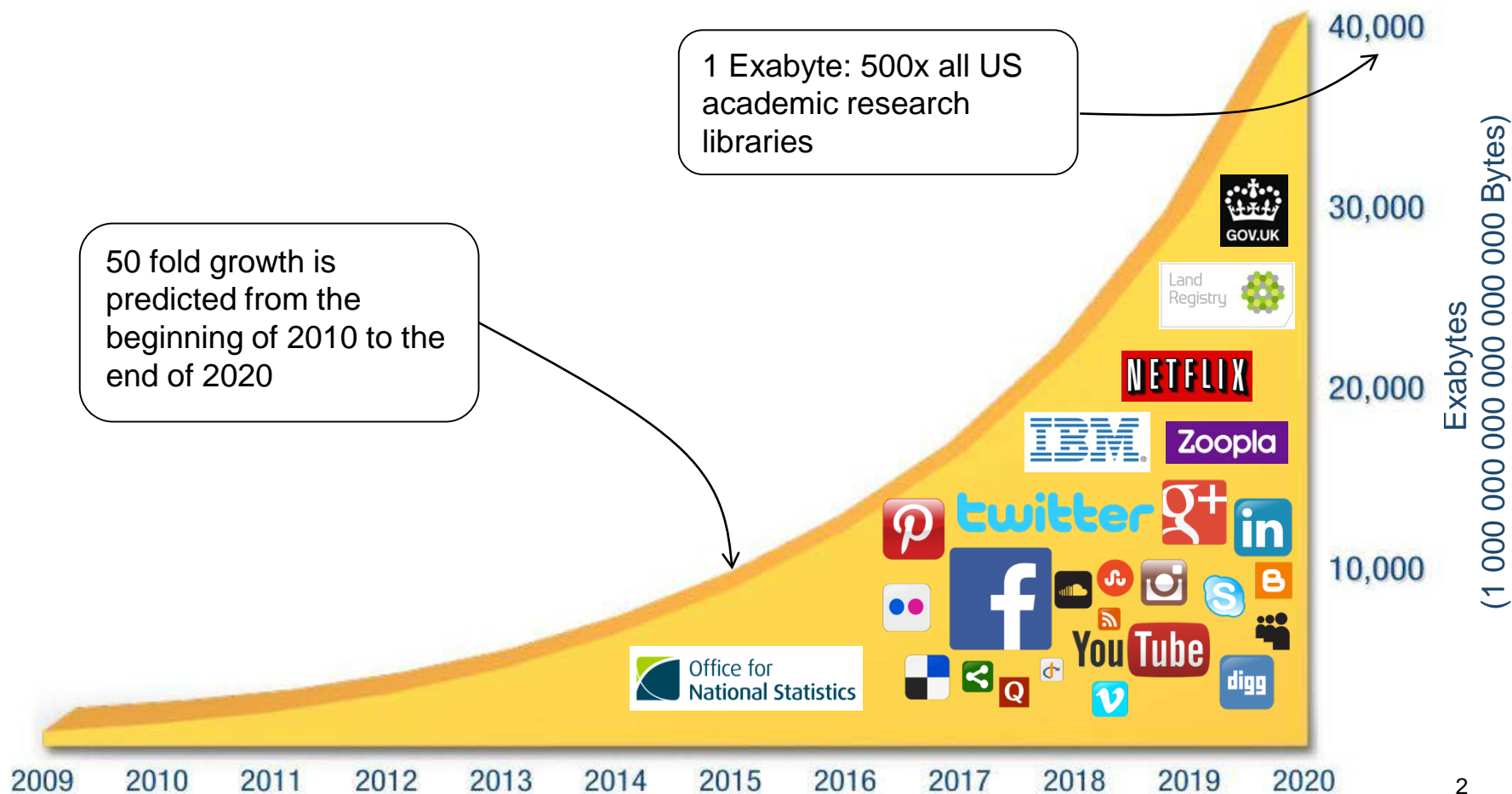
29th June 2017

Stephen Aldridge
Director, Analysis and Data
Department for Communities and Local Government

There has been an exponential increase in the volume of data available, from private and public sources, as well as social media
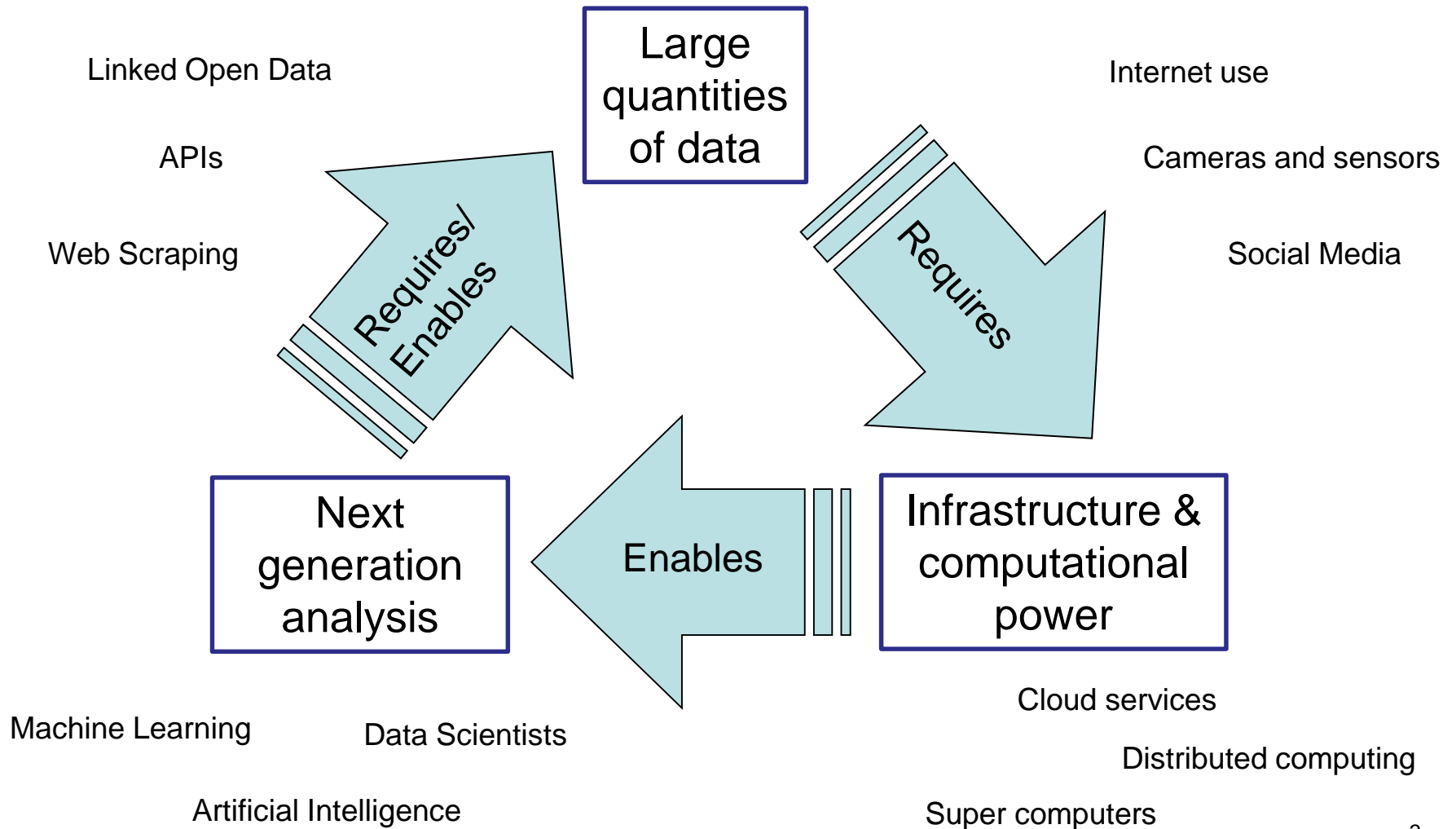
It is estimated that almost **90% of the world's data was produced during the last 5 years**

1 Exabyte: 500x all US academic research libraries

40,000

30,000

50 fold growth is predicted from the beginning of 2010 to the end of 2020

20,000

10,000

Exabytes
(1 000 000 000 000 000 000 Bytes)

2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020

2

Source: IDC's Digital Universe Study, sponsored by EMC December 2012

Department for Communities and Local Government

This comes hand in hand with an explosion in different ways of collecting, accessing and analysing data

Linked Open Data

APIs

Web Scraping

Large quantities of data

Internet use

Cameras and sensors

Social Media

Requires/ Enables

Requires

Next generation analysis

Enables

Infrastructure & computational power

Machine Learning

Data Scientists

Artificial Intelligence

Cloud services

Distributed computing

Super computers

Effective use of this data can transform the public sector and public services

Using data effectively can transform public services by:

- Improving **understanding of current problems** and **predictions** of the future;

- Providing better **evidence** of what does and does not work;

- Helping to better target interventions at the right people- improving outcomes and reducing costs

- Driving **innovation and diffusing new technologies across the public sector**; and,

- Improving **accountability** to citizens.

airbnb — Uses data to predict the optimum price for a host to rent their property.

T··Mobile — Combined transaction data with information from social media to halve customer defections

A US NGO (Polaris) created a mapping tool to identify human trafficking networks. This has now been adopted for law enforcement in several states.

Birmingham City Council is using air temperature sensors to target gritting vans to roads covered in ice and snow while avoiding unaffected roads.

Department for
Communities and
Local Government

First, by **consolidating systems** to collect and manage statistical and administrative data – including the collection of all our data online
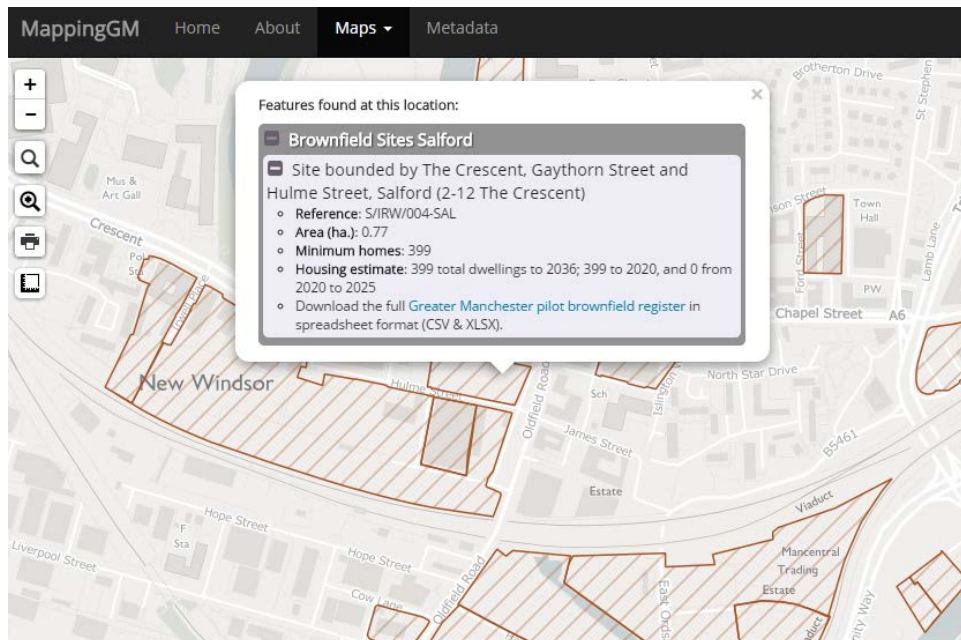
Improves the data provider experience

One central secure and accessible location

Greater efficiency & resilience in our data collection teams

Why?

Frees up resources for analysis

Further improves data quality

More robust and reliable system

DCLG has been transforming the way it collects, uses and disseminates data

Second, by **developing standards** to facilitate the unlocking and linking of data

Developed data standards for the new local brownfield land registers and used these to gather and consolidate local registers from around the country.

Collaborated with the Government Digital Service to build the single, definitive register of Local Authorities in England.

Department for
Communities and
Local Government

Third, by making a reality of **open data** through the creation of an open data platform (OpenDataCommunities) – a Whitehall leader

- **Over 200 datasets** in open and linkable formats
- Data from all **15 million Energy Performance Certificates** (EPCs)
- An **interactive mapping tool** to explore 2015 indices of multiple deprivation

Benefits:
- **Internally** - to power briefing tools, to help us shape and understand policy outcomes
- **Externally** - to improve collaboration and decision-making based on the facts.
- **Linking -** to re-use DCLG sources alongside others' own and other 3rd party data

OpenDataCommunities
Open access to local data

Explore ▾   Tools ▾   Search ▾   Help

The official UK Department for Communities and Local Government linked data platform.

Council Tax Band D Average: a data cube spreadsheet
This dataset shows the average Band D council tax amount charged by each billing authority and county council in England. Billing authorities comprise shire districts, unitary authorities, metropolitan districts and London boroughs.
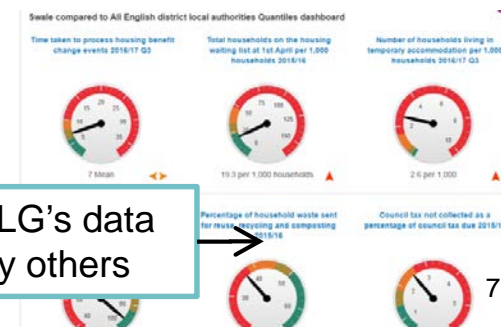
DATA   ABOUT   API

Spreadsheet: Average Council Tax

Filter **Areas** by type:
Any (no filter)

Reference Area	2015-2016	2010
Countries		
England E92000001	1,483.58	
English Upper Tier Authori...		
Bath and North East Some...	1,479.22	
Bedford E06000055	1,577.11	
Blackburn with Darwen E0...	1,495.41	

Map
The data have been sorted into coloured groups, each of a similar size. Hover over any area for details.

Anyone can explore, visualise and extract our data in-situ….

Allowing DCLG's data to be used by others

7

Using our open data platform we have:

- Worked with the Local Government Association automatically to feed a selection of DCLG datasets straight into their websites

- Worked with individual local authorities to associate local data with related DCLG sources

- Collaborated with Shelter and Centrepoint UK to provide housing and homelessness data over the wire for re-use alongside the charities' own local data and intelligence

- Developed inter-active and constantly updated policy performance dashboards for Ministers and senior officials
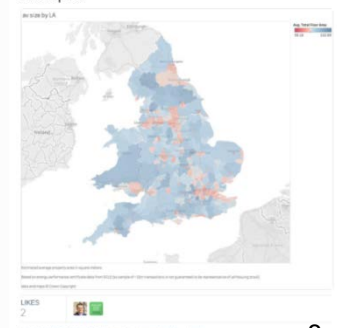
Using our open data platform.....
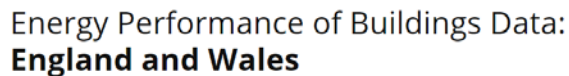
others have rapidly generated analysis of their own – as powerfully illustrated by the release of EPC data

Department for Communities and Local Government

Fourth, DCLG has **linked administrative personal data** on a substantial scale for the first time

To evaluate the troubled families programme:

- We have linked data from **over 150 local authorities** with outcomes ranging from crime and ASB to education and school attendance to child safeguarding to financial exclusion and work to health and to domestic abuse or violence

- This will enable us to analyse what has happened to families and individuals up to 5 years before and up to 5 years after the intervention

- We will also be able to compare the intervention group with a comparison group

- We currently have linked data on around **350 000 individuals**

Linked admin information includes:

Police National Computer

National Pupil Database

Hospital Episode Statistics

Work & Pensions Longitudinal Study

Department for
Communities and
Local Government

We are now building on the success of linking administrative data for the troubled families evaluation, applying it to homelessness

Currently underway are a major redesign of homelessness statistics and two policy evaluations.

In line with new homelessness prevention legislation, we are overhauling our statutory homelessness statistics. We will collect **case-level data** from councils on the at-risk and homeless population.

This will allow a greater understanding of risk factors, what works to prevent homelessness, and the size and characteristics of this cohort.

We are intending to link this to other administrative data, to understand the links between homelessness and a broad range of cross– government issues.

Department for
Communities and
Local Government

We have learnt the **importance** of:

- data standards for consolidating and linking data.
- legal or perceived legal barriers.
- upgrading data science skills across both the analytical professions and the policy profession in the Civil Service.

We have learnt there are **costs**:

- in preparing and releasing of data as open data.
- in putting in place privacy or other safeguards.
- in recruiting and retaining key staff

We have learnt there are **constraints**:

- in accessing/linking health data.
- on sharing even anonymised central government administrative data with for example with local authorities or other sub-national bodies
- on who can work on anonymised linked datasets – depending on the agreements that underpin the data linking.

Department for Communities and Local Government

The Government has committed to:

Publish much more information about public services online to:
- Hold public services to account
- Help the public make decisions (such as 'schools maps' providing key information for parents choosing schools)

Continue the **drive for open data**

Create a **comprehensive geospatial data body** within government, which will be the largest repository of open land data in the world.

The future is bright if we are bold

There is huge potential for the future.

As vast quantities of data combine with digital advances.

A wave of change is to come.

We have developed an prototype area profiler which automatically generates briefing for Ministers drawing on data transferred over the wire using OpenDataCommunities and similar sources, and algorithms to upload the right data on the topic specified by the user.

Is this embryonic artificial intelligence? The first step towards robots briefing Ministers?

Today's Civil Servant

Tomorrow's Civil Servant?