# Experience, Business Models and Case Study of Open Data in UK

Andrew Stott

UK Transparency Board
*formerly* Director, *data.gov.uk*
& UK Deputy GCIO

@dirdigeng
andrew.stott@dirdigeng.com
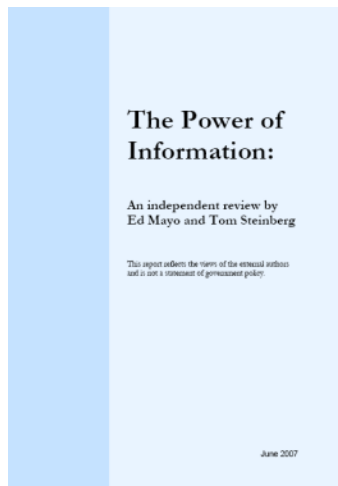
Taipei
18 Jan 2012

# Presentation Contents

- Drivers of UK Open Data Policy

- Business Models for Open Data

- Ecosystem for reuse of Open Data

- How data.gov.uk was developed

- Lessons Learned

- Unforeseen additional benefit
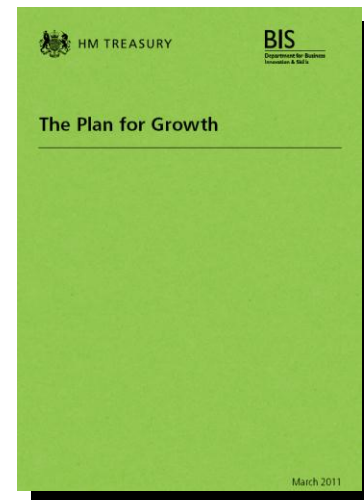
# Drivers of UK Open Data Policy

# Economic Growth and Social Value



June 07

Feb 09

Mar 11

New economic and social value

# New Businesses

# Social value and cohesion
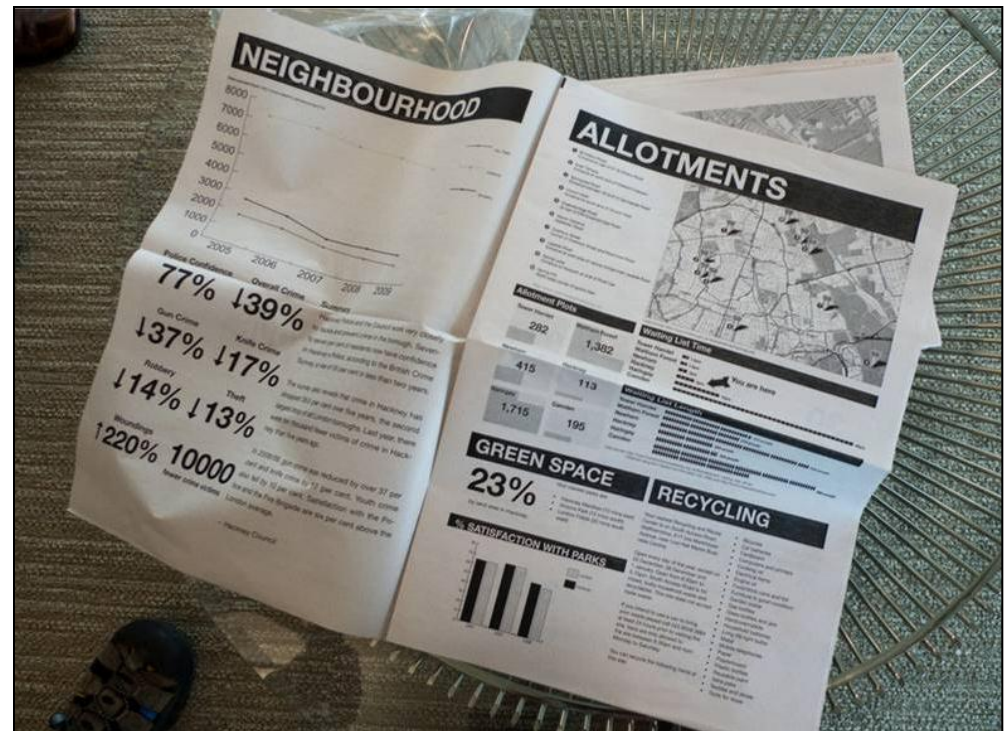
# Improve Public Services & Enhance Choice
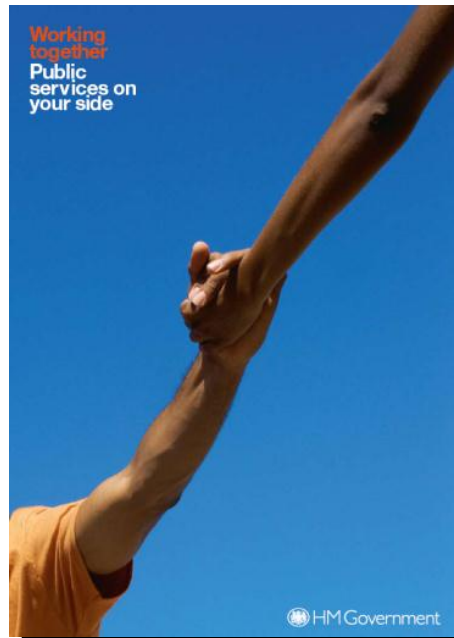


**Mar 09**
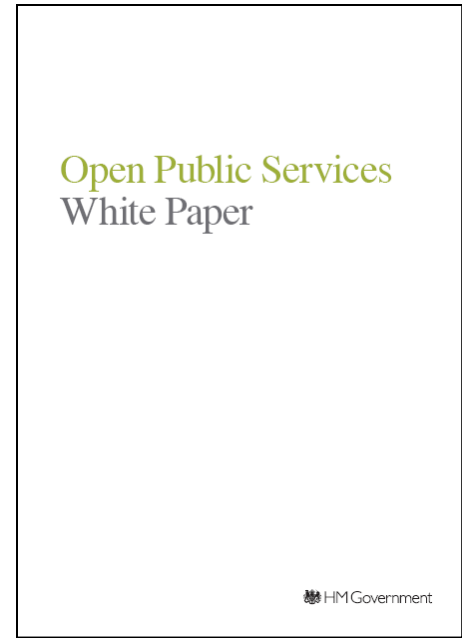
**Jul 11**

**Improve public services**

**New economic and social value**

# Benefits: Improving Public Services

- Gives "local" view of services
- Gives "holistic" view of all local services
- Allows people to give feedback, lobby for improvement, participate in policy-making
- New apps reach people who do not visit government websites
- Information presented in more relevant ways
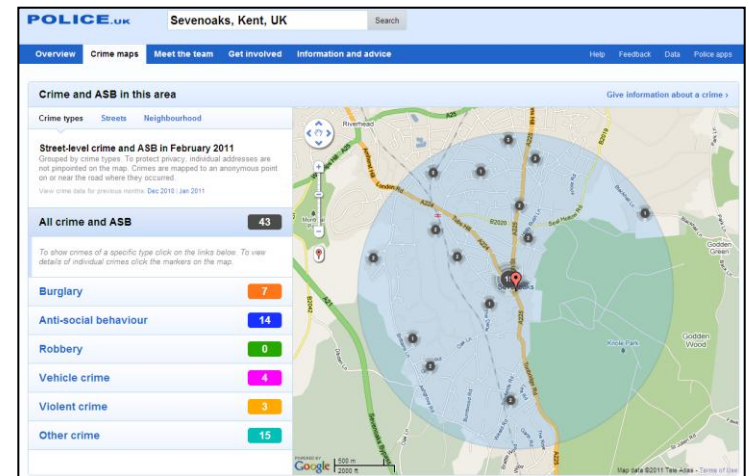
# Transparent and Accountable Government



**Jun 09**     **May 10**

**Transparent & Accountable Government**

**Improve public services**

**New economic and social value**

# Benefits: Accountable Government

- Transparency allows citizens to hold government to account for what they do
- Strengthens pressures to comply with ethics codes - decisions and spending are more public
- Puts pressures on suppliers to be fair in dealing with government
- Politicians judging that it's more important to be open than to avoid criticism
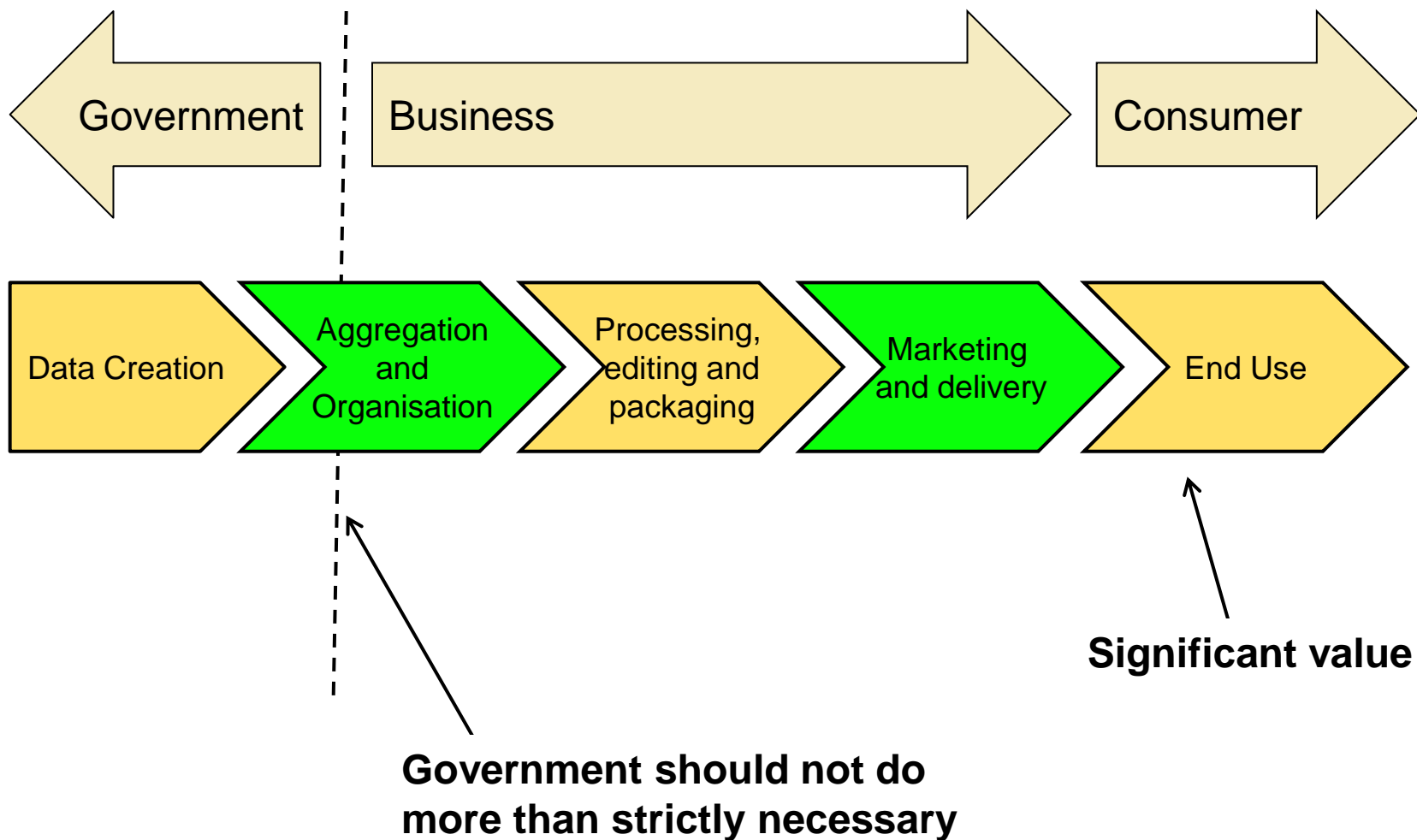
# Business Models for Open Data

# Evidence for EU PSI review

- EU27 direct gains from completely open access to PSI could be ~ €40 Bn

- Aggregate economic impacts could be even larger: ~€140 Bn

- Direct market associated with PSI use less important than spillovers and new uses.

- Future innovations due to easier access add further economic and social benefits

- Costs and revenue foregone? ~ € 1 bn

Source: Graham Vickery, Information Economics

# Public Data Re-Use Value Chain

Government

Business

Consumer

Data Creation

Aggregation and Organisation

Processing, editing and packaging

Marketing and delivery

End Use

**Government should not do more than strictly necessary**

**Significant value**

# Emerging Business Models

# Cleansing and organising data



Business Intelligence from item-level purchasing data

15

# Information services to the public



Weather, transport and public facilities among most downloaded Smartphone Apps

# Data Mining



PharmaTimes ONLINE

News   Magazine   JobSearch   Competitions   Events/Meetings   Gallery

WORLD NEWS | UK NEWS | CLINICAL NEWS | FREE NEWS HEADLINES | FREE RSS FEED

Where am I? >   Article

## Govt open data policy "set to boost UK life sciences growth"

UK NEWS / WORLD NEWS | DECEMBER 01, 2011    SHARE  +1  1

LYNNE TAYLOR    Tweet  11

The UK government's Autumn Statement includes "world-leading commitments" to boost growth in the nation's life sciences sector by transforming access to public-sector health and care data.

From next September, services provided by the NHS Information Centre for Health and Social Care will link datasets from primary and secondary care, which will enable professionals in the NHS, pharmaceutical industry, academia and elsewhere to access unprecedented and unequalled levels of information about the journeys of patients through the care system and the outcomes of different treatments, it says.

- Prescription data
- Patient outcome data
- Longitudinal health records
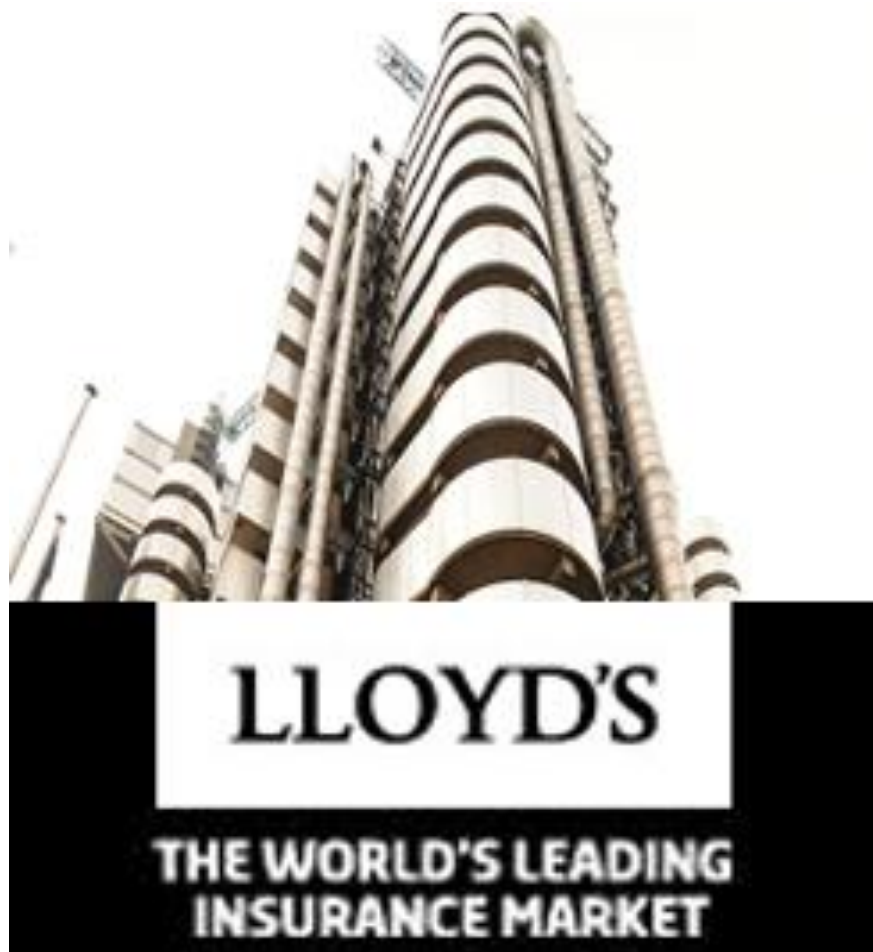- Pupil-level education records

17

# Operational efficiency/optimisation
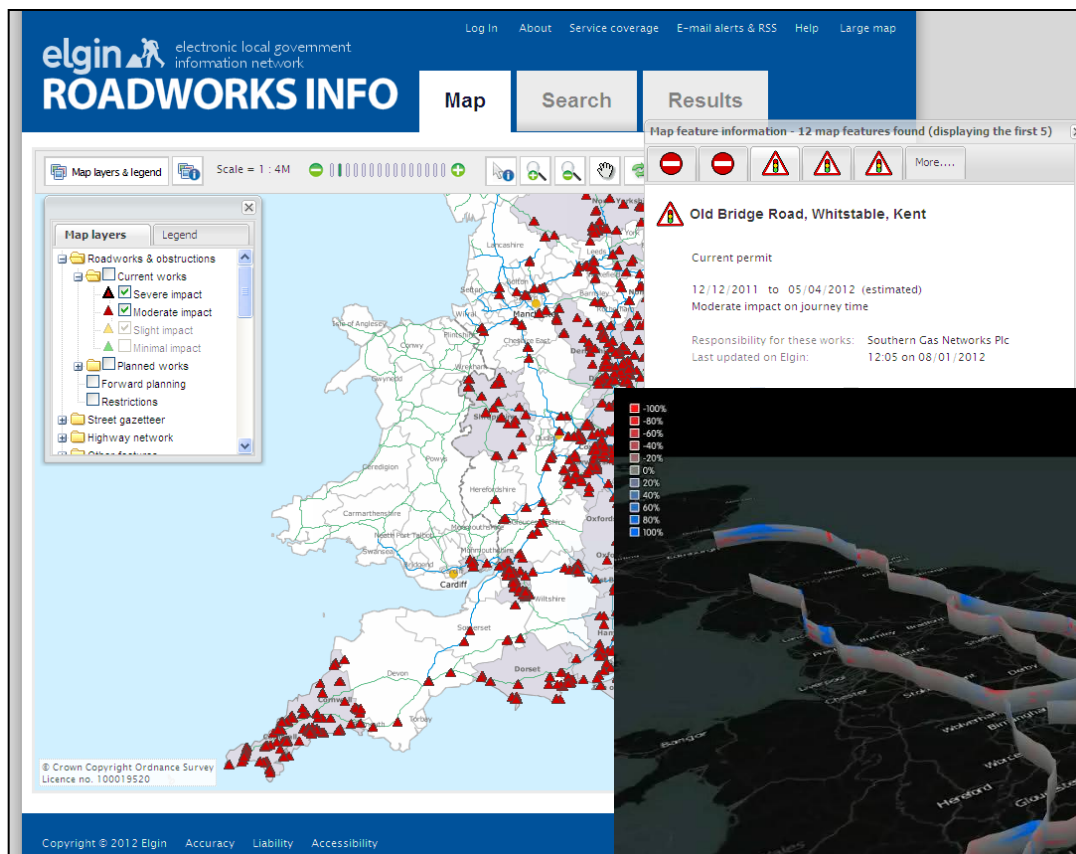


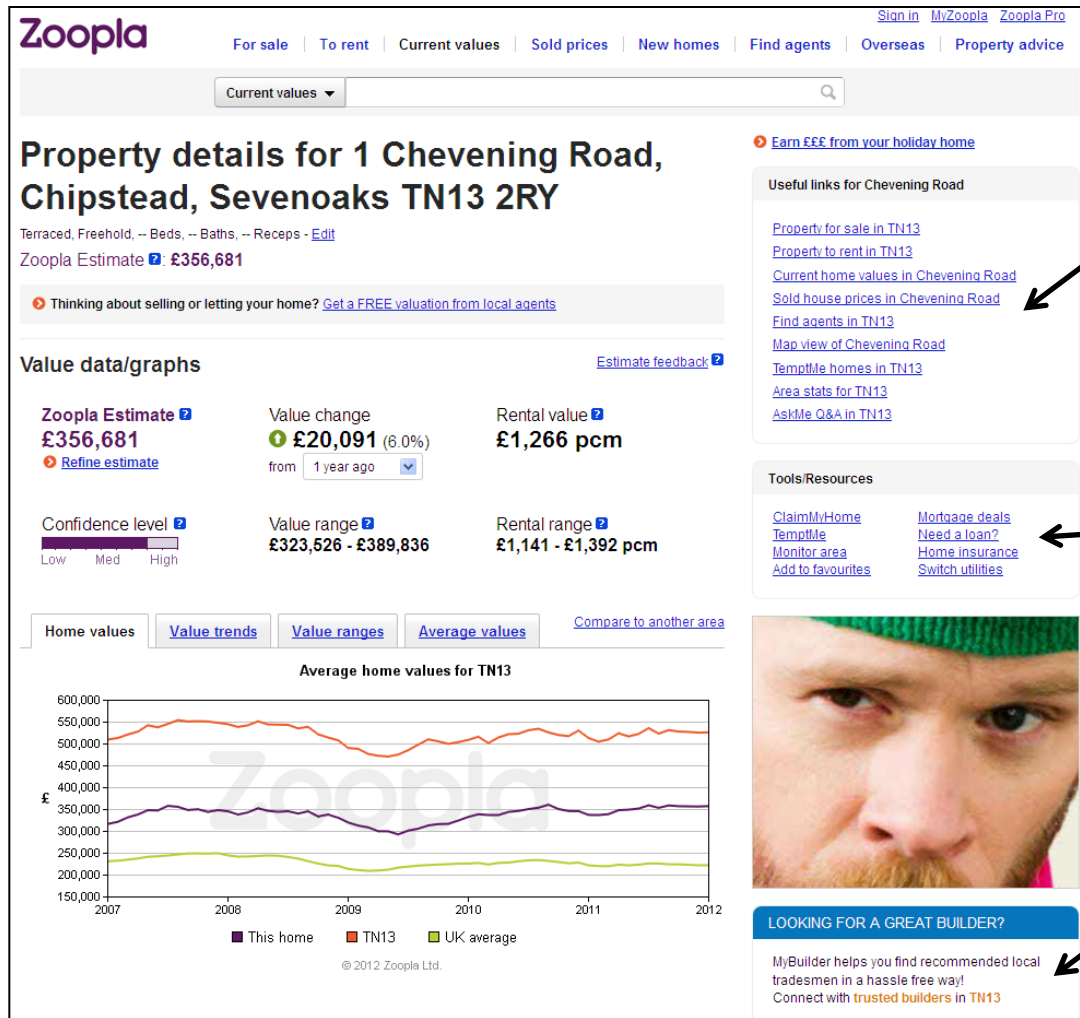Real time info on road delays and roadworks allows logistics efficiency

# Financial Products



LLOYD'S

THE WORLD'S LEADING INSURANCE MARKET

- Weather Risk Management (US: $4bn annual contracts value)

- Flood insurance on detailed topography and river records

# Customer attraction and retention



estate agents/ realtors

Financial services

builders and other local services

Local house prices attract potential customers

# Data Broking



Customer-focused reliable high-volume serving of data and APIs

# Emerging Business Models: Summary

Models already seen:

- Cleansing and organising data
- Information services to the public
- Data mining
- Operational efficiency/optimisation
- Financial products
- Customer attraction and retention
- Data broking

# More Innovation



Opening up mapping data and providing a developer platform has increased innovation

# Public Data Re-Use Value Chain (first movers)



Government

Business

Consumer

| | | | |
|---|---|---|---|
| Aggregation and Organisation | Processing, editing and packaging | Marketing and delivery | End Use |

Data Creation

| | | | |
|---|---|---|---|
| Aggregation and Organisation | Processing, editing and packaging | Marketing and delivery | End Use |

| | | | |
|---|---|---|---|
| Aggregation and Organisation | Processing, editing and packaging | Marketing and delivery | End Use |
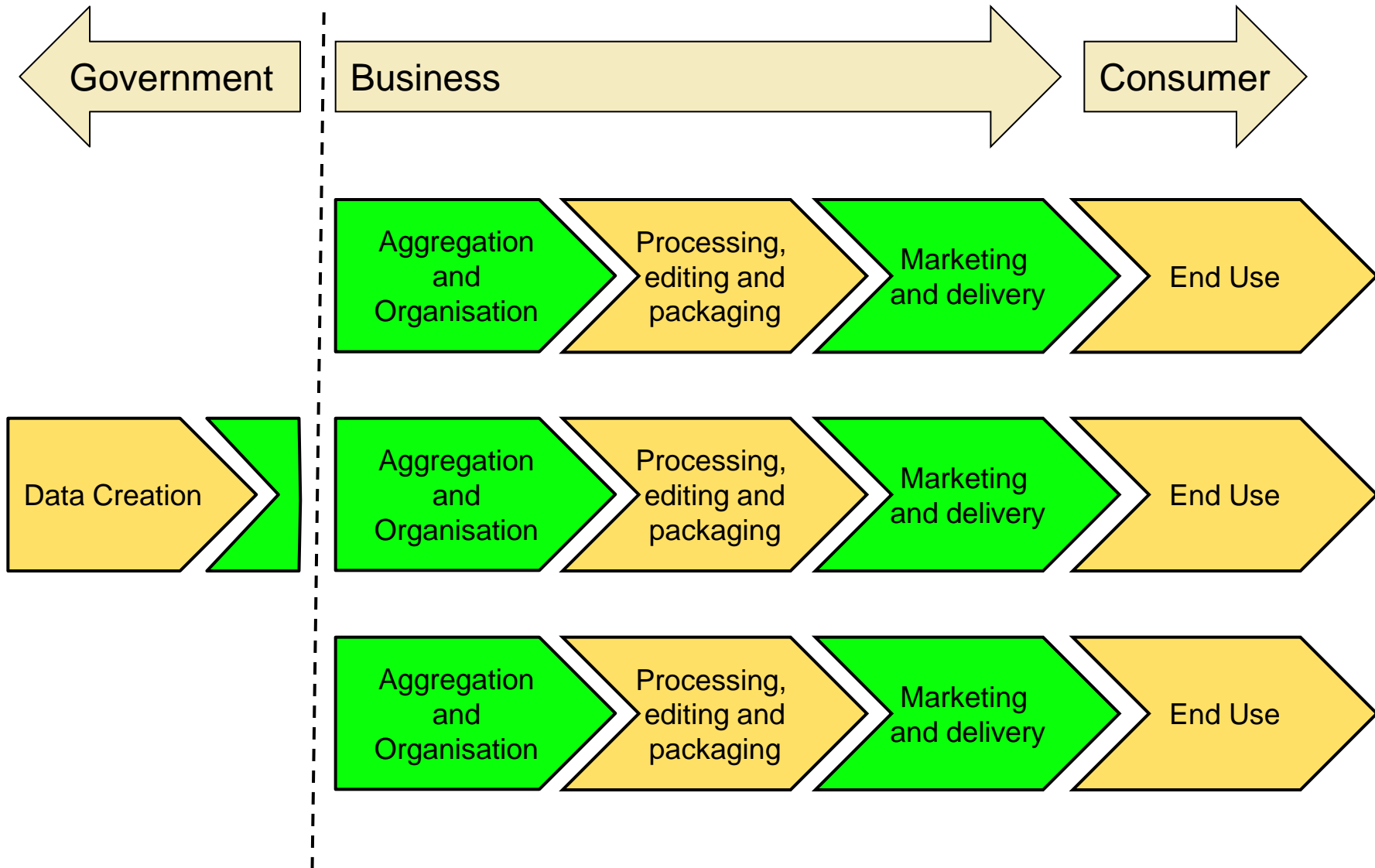
# Public Data Re-Use Value Chain (maturing)
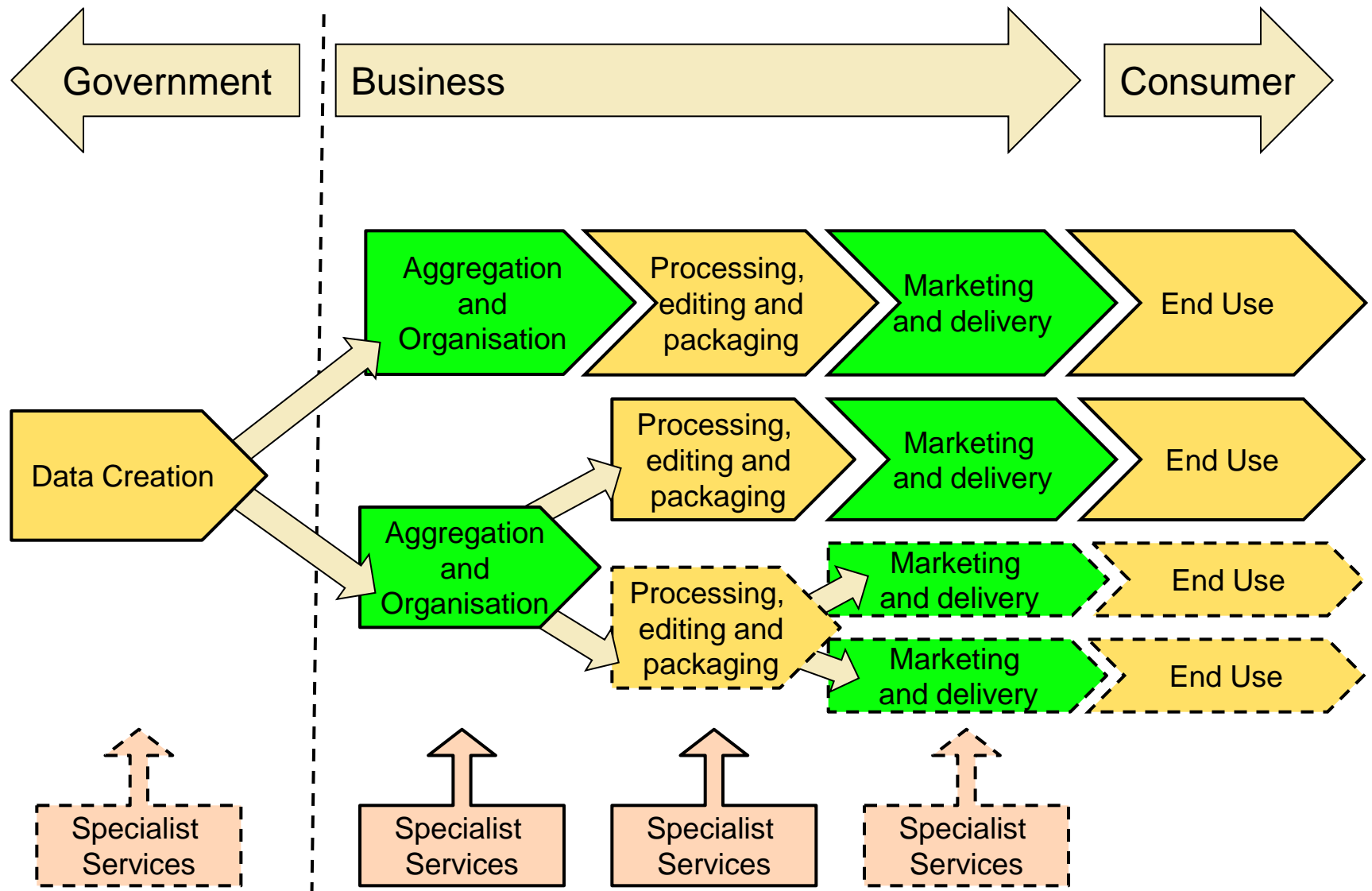
# Ecosystem for business re-use of Public Data

- Data accessible through a single catalogue
- Predictable pricing
- Open Licensing
  - No restriction on innovation
- Common, open formats and ontologies
- Availability of skills and specialist services
  - Open Data Institute, LinkedGov, Semantic Web
- Data Broking and other dis-aggregation of value chain

## It's not just about new data

Scope for "Open Data" also includes data previously "published" but …

- in non-reusable format
- with restricted licence
- only aimed at specialist groups
- only for payment
- only in response to requests
- difficult to find

data.gov.uk contains a lot of data which nobody knew was already published

27

# Charging?

**Free is the default**

- Taxpayer has already paid for most data
- Economic gain maximised when priced at marginal cost of distribution
  - Raw data free
  - Possible small service charge for reliable API
- Sales revenue not material to most agencies
- Sales administration typically >25% of revenue

**Exception: where data <u>collected</u> for sale**

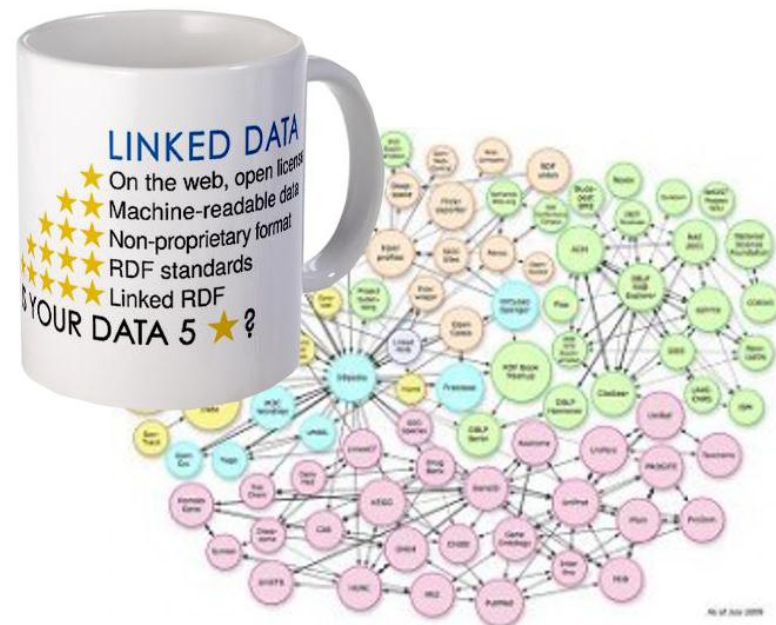- "Public task" data free, "value added" charged

# Clear, common, licensing approach

- "Open" licence essential for unrestricted re-use
- Clarity vital, especially for businesses
- Common licensing for all Government data valuable
- UK OGL
  - free to reuse
  - with attribution
  - without misrepresentation
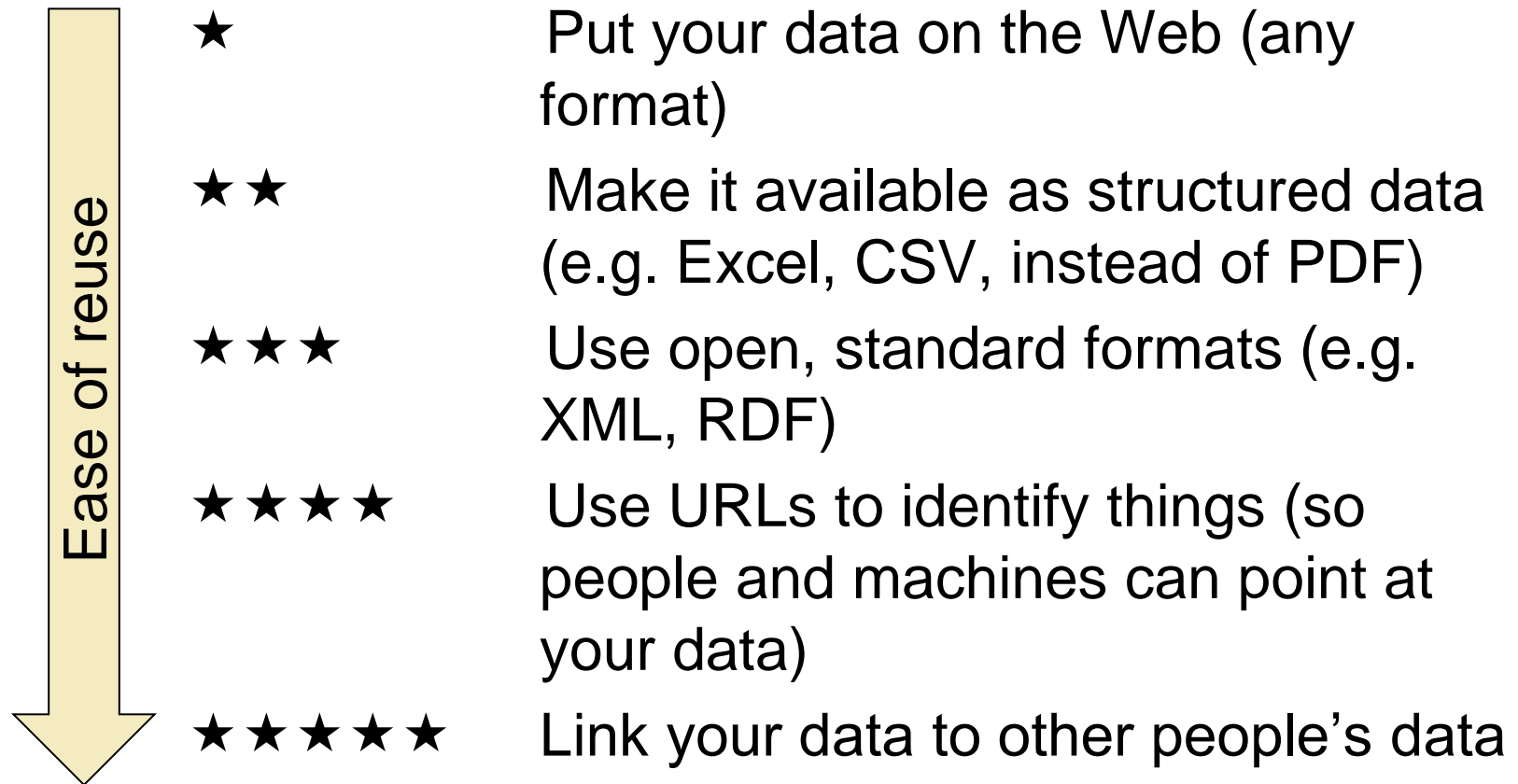- CC-BY a good alternative
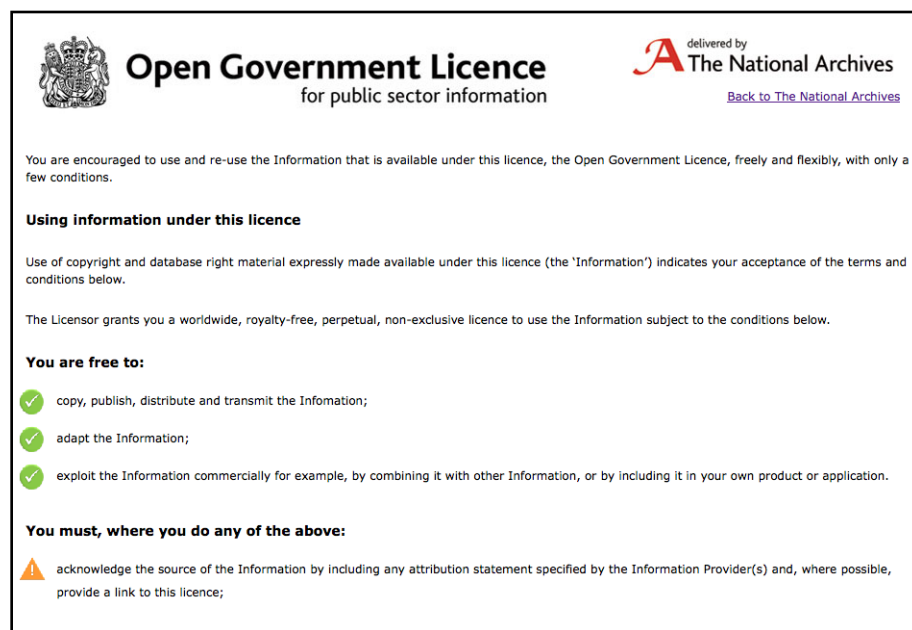
# Government: key provider of reference data

- Government Registers/Licences are key reference data for wider economy – the "National Information Infrastructure"

- To maximise economic benefit the data needs to be easily reusable

- Today that means "linked data" – the building blocks of the Semantic Web



LINKED DATA
On the web, open license
Machine-readable data
Non-proprietary format
RDF standards
Linked RDF
IS YOUR DATA 5 ⭐ ?

# Data Publishing – Star Quality

Ease of reuse →

★        Put your data on the Web (any format)

★★      Make it available as structured data (e.g. Excel, CSV, instead of PDF)

★★★    Use open, standard formats (e.g. XML, RDF)

★★★★   Use URLs to identify things (so people and machines can point at your data)

★★★★★ Link your data to other people's data

# Continuously engage with developers









*Photos: @memespring,*
*@MadLabUK, @paul_clarke*

# .. and highlight applications, not data

# Open Data Institute: its mission

- Develop capability of UK businesses to exploit value of Open Data

- Engage developers/small businesses to build Open Data supply chains and commercial outlets

- Help public sector use its own data more effectively

- Ensure academic research in Open Data technologies

# How data.gov.uk was developed

# data.gov.uk organising principles

- Small central team (4 people)
- Short stages, each with clear goals
- Respond quickly to feedback from users
- Draw in resources from Government's experts
- Each dataset always controlled by its owner
- data.gov.uk holds metadata not data!
- It is not an IT project! Keep the IT simple
  - use Open Source (CKAN, Drupal, etc)
  - use existing contracts/infrastructure with niche firms
  - host data on existing websites or on public Cloud

# The first two years of data.gov.uk
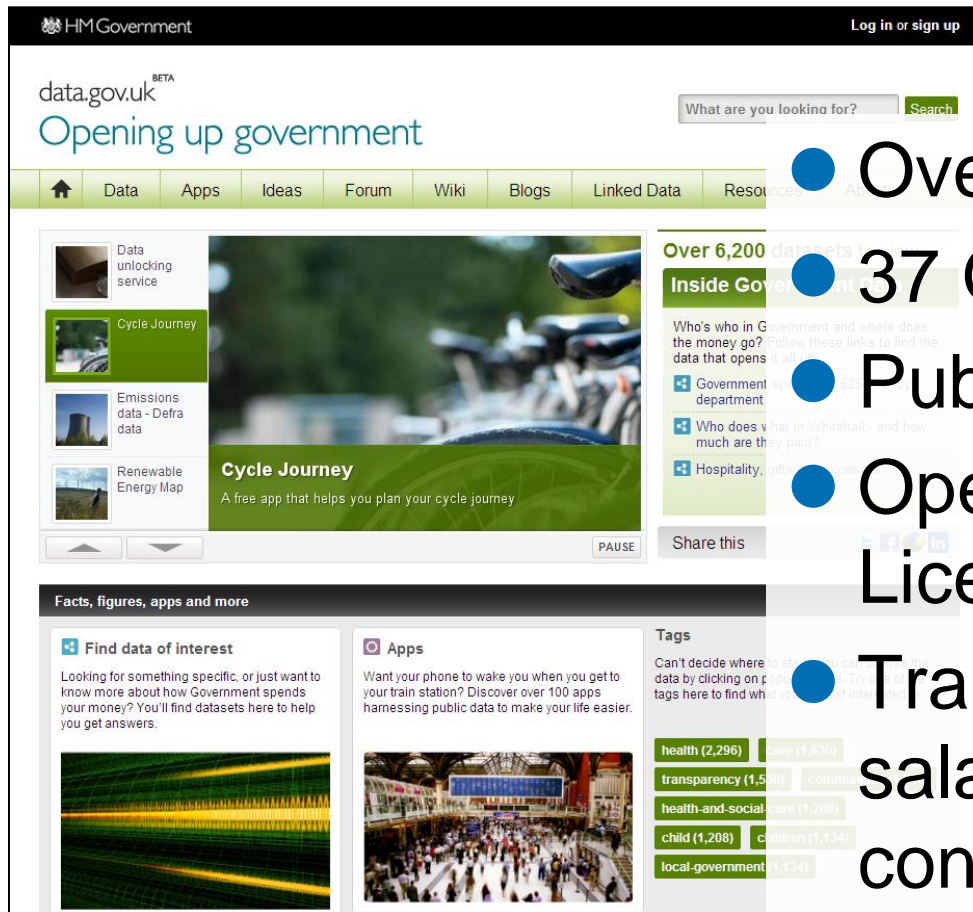


- Over 7500 datasets
- 37 GB of geo data
- Public Data Principles
- Open Government Licence
- Transparency of salaries, spending, contracts and tenders

# Lessons learned
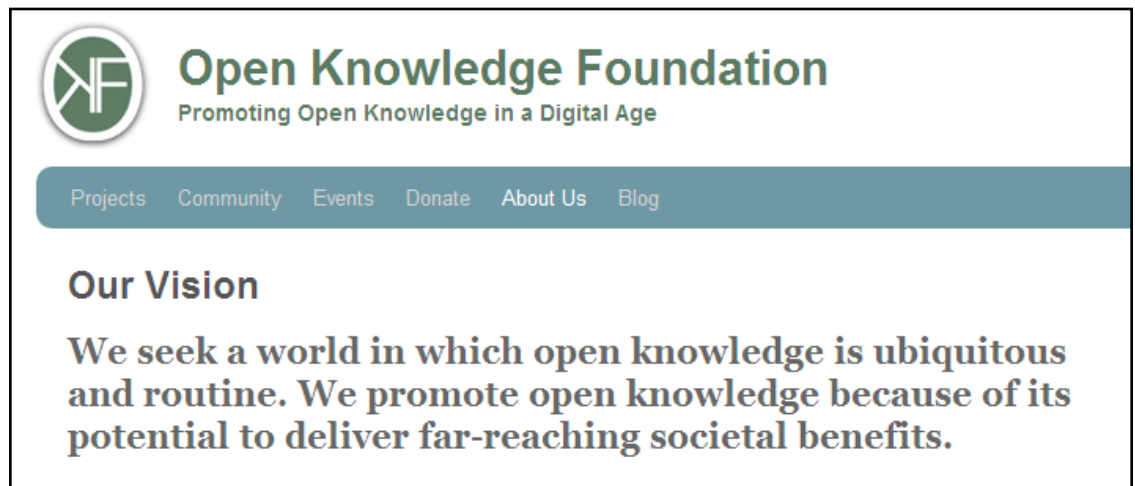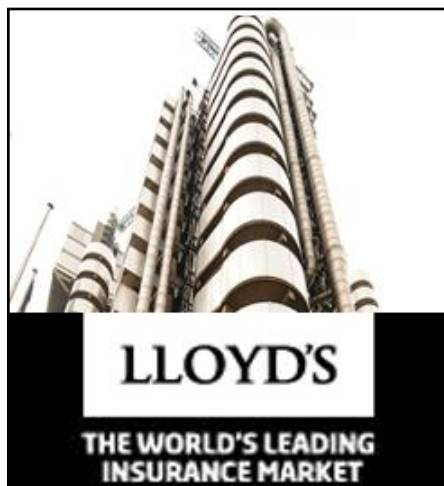
# Important to have top-level political support





"Greater transparency across Government will enable the public to hold politicians and public bodies to account"

"Public information does not belong to Government, it belongs to the public on whose behalf government is conducted."

# Important to have strong "demand-side"

# Also important  to have passionate team!

# Refine policy principles based on experience



**http://data.gov.uk/blog/new-public-sector-transparency-board-and-public-data-transparency-principles**

# Incremental delivery

- Allows "data pipeline" to be turned into automated tools and repeatable processes
- Enables future releases to be prioritised by developer feedback
- Creates steady "news flow"
- Can postpone difficult cases

### Datasets on data.gov.uk
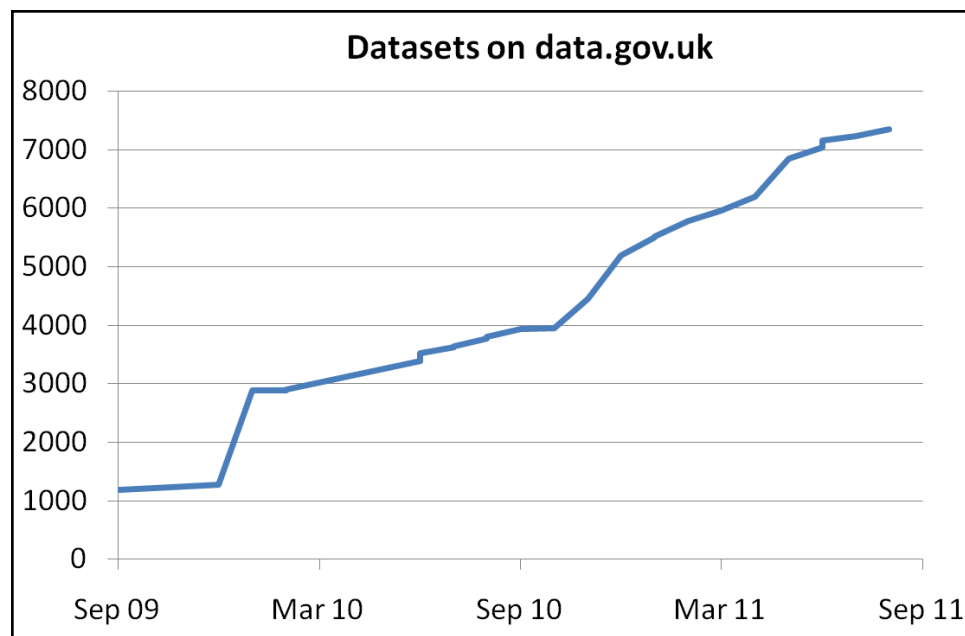
## Handling the concerns of data owners


TIM**BERNERS-LEE**
(click to play)

- Releasing data is a new idea
- Look for ways to manage perceived risks
- Look for examples and precedents
- For first phase, be prepared to compromise on data content – but not on licence terms
- "Hackdays" can show potential benefits
- Ensure data owners get credit for what they do release

# Manage expectations, prepare for mistakes



"We're making a small start next week. But eventually, it's going to make a big difference."

"The information we're publishing next week won't be perfect, it won't always be in the most convenient format, and I'm sure there'll be some mistakes. But I want to get on with it."

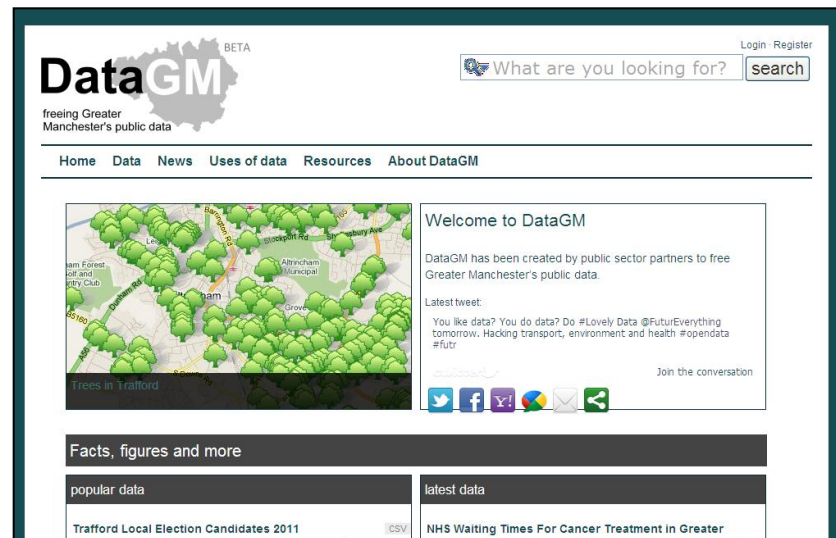*UK Prime Minister 29 May 2010*

# … and the biggest lesson of all

**Overcome obstacles
practically
by doing,
not debating**

# Unforeseen additional benefit

# Government is a data user too

- Once Open Data is published, other public agencies can use it too

- Each level of government can see all the data about an area from all the different levels

- Greater Manchester estimated potential savings of £8.5m a year from time saved in looking for data

# Questions?

# End