Data and Transport

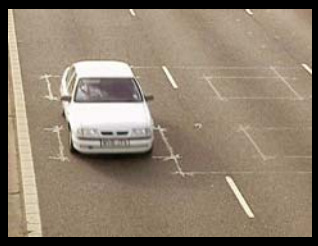
Professor Frank Kelly FRS
Faculty of Mathematics, University of Cambridge
Chief Scientific Adviser, Department for Transport

Foundation for Science and Technology
9 May 2006

Examples of transport data

- MIDAS (motorway incident detection and automatic signalling)
- Transport Direct
- Accession
- Speed Limit database
- New sources of transport data

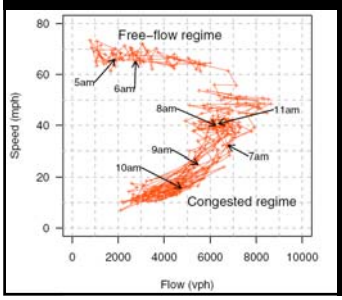
MIDAS

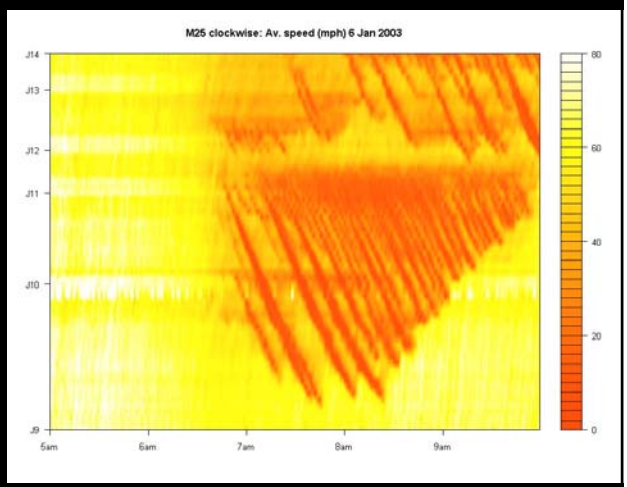


Real-time control of speed-limit signs.
Data archived since 1997.

www.ha-research.gov.uk

Flow breakdown





Data gold mining: MIDAS and journey time predictors. Richard Gibbens and Wiebke Werft, Significance 2, September 2005

MIDAS

- Initial purpose was real-time closed loop control of speed limits
- Archived data available for analysis, and able to provide insight into complex system behaviour
- Or, be fused with other data for applications not possible to envisage at time of data collection:

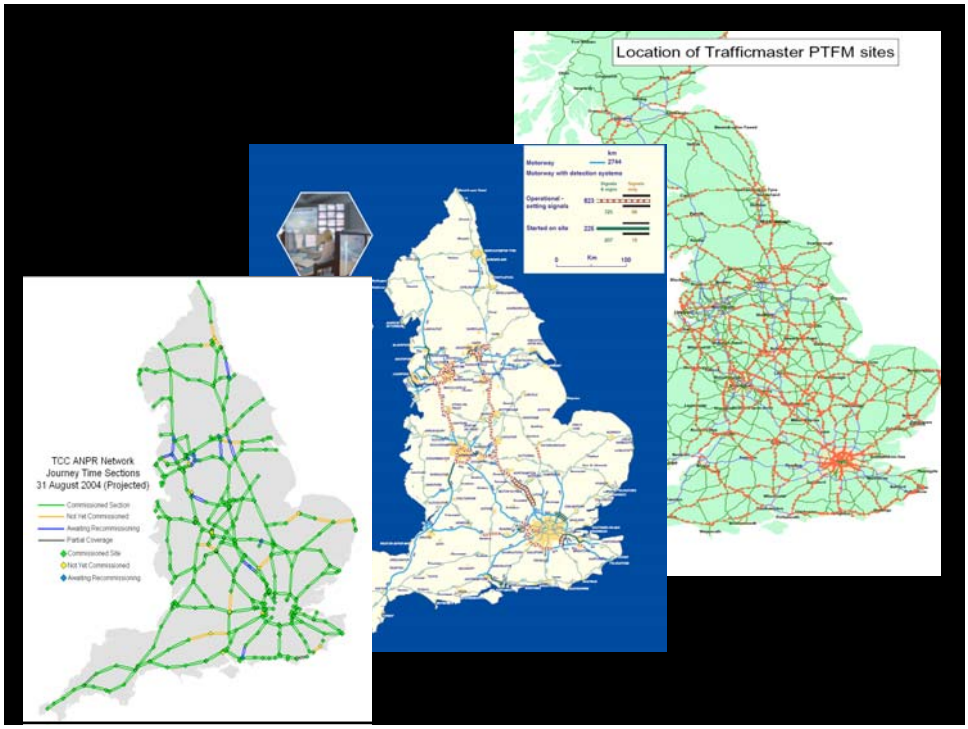
PSA target: reliable journeys

Target measured by **average vehicle delay** on each of 98 routes along the strategic road network.

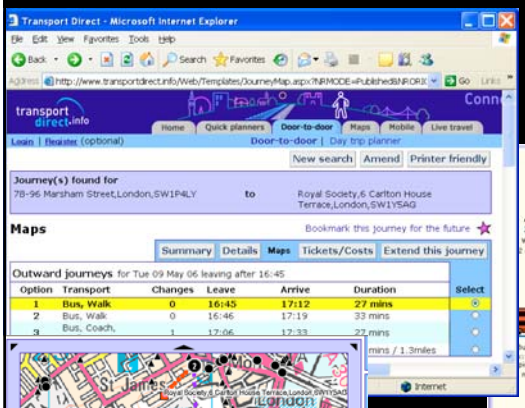
- **Average delay** (time loss per vehicle km) is the difference between observed journey time and JT at a reference speed :

  = ((*JT – refJT) * flow) / total vehicle kilometres*

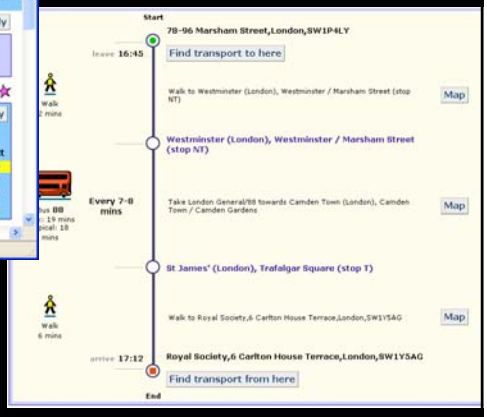
- The target will be achieved if the average vehicle delay on the 10% slowest journeys is less in 2007-08 than in the baseline period.
- Target requires integration of data from a variety of sources: MIDAS, Trafficmaster, NTCC, ITIS.



Transport Direct

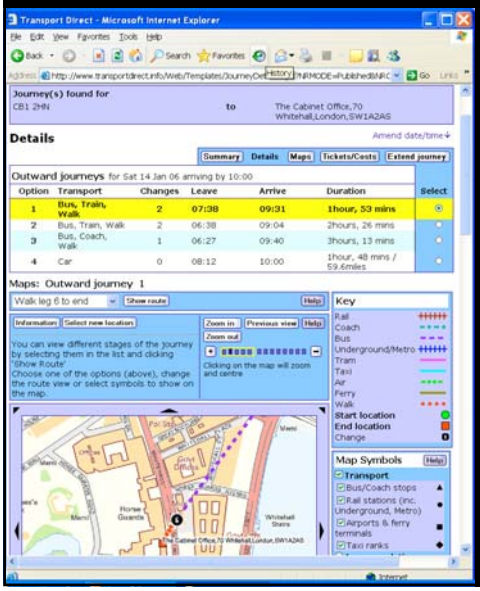


The vision for Transport Direct is to provide a comprehensive, easy-to-use multi-modal travel information and ticketing service



www.transportdirect.info

Transport Direct



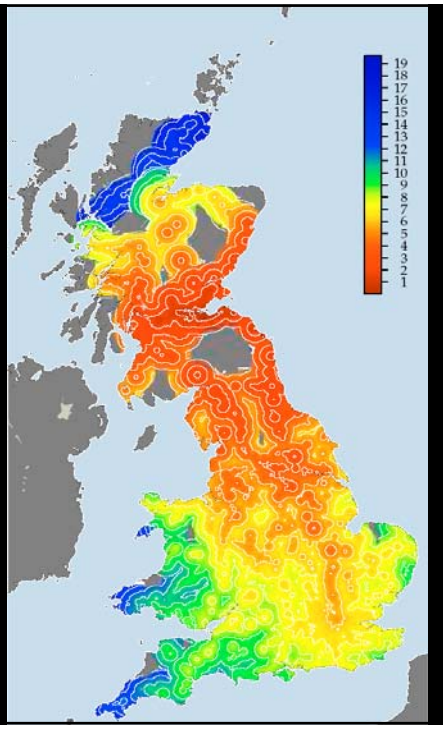
Transport Direct is, in reality, an enormous virtual team incorporating hundreds of organisations and individuals
Nick Illsley, Chief Executive

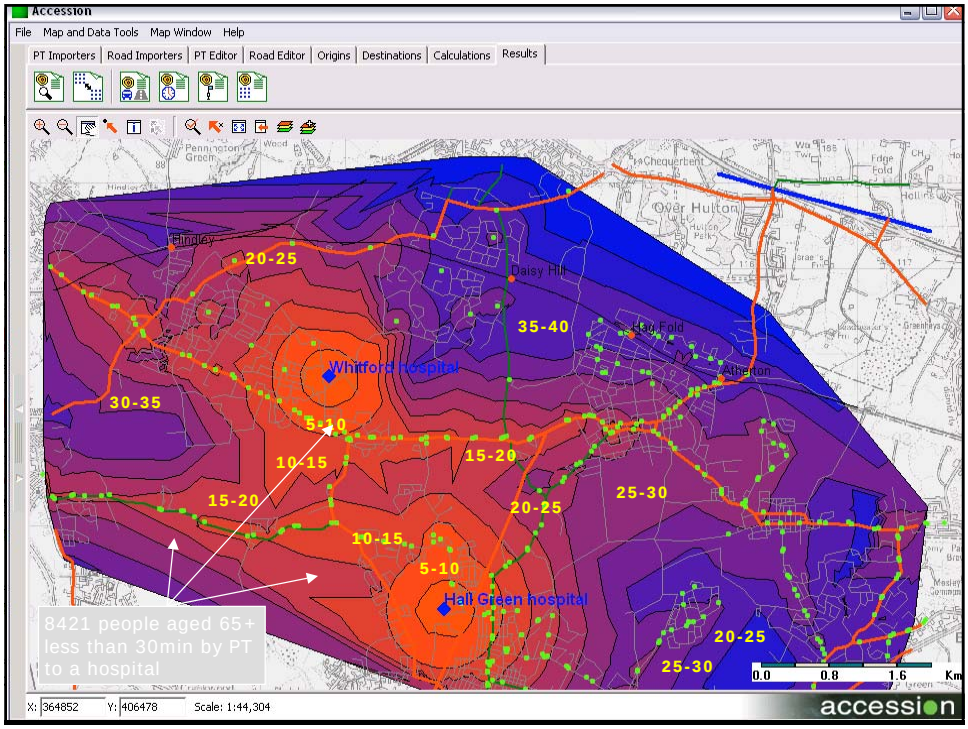
Each of over a hundred sources needed to agree to provide their data and also to make it available in common standards and formats. (Example task: agree a unique number for each bus stop in the country…)

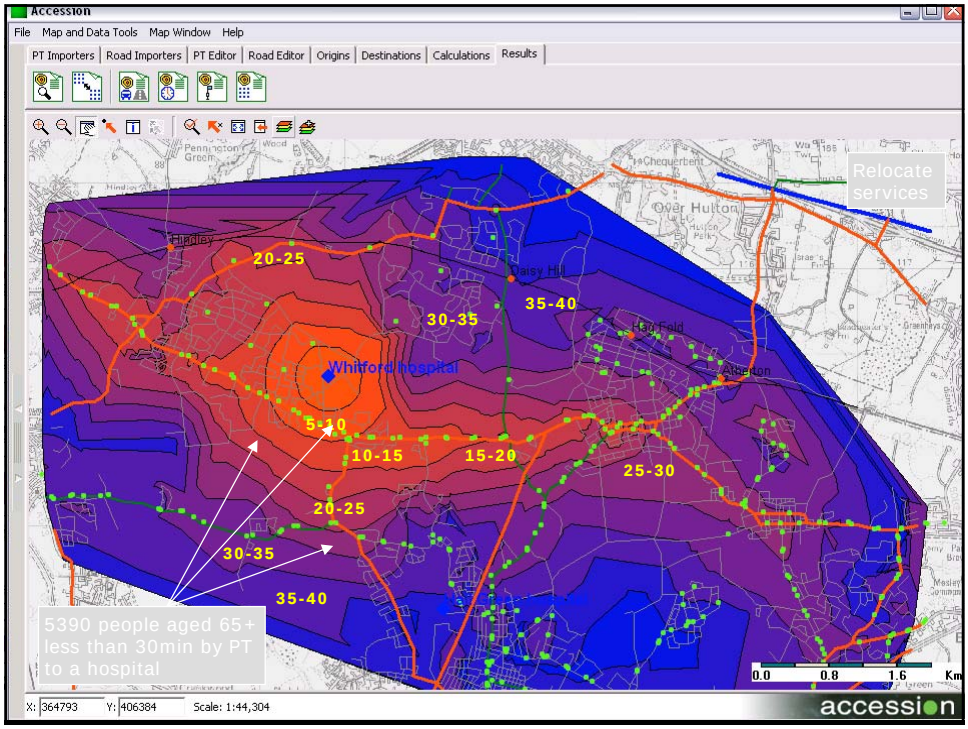
Data useful for many other applications….

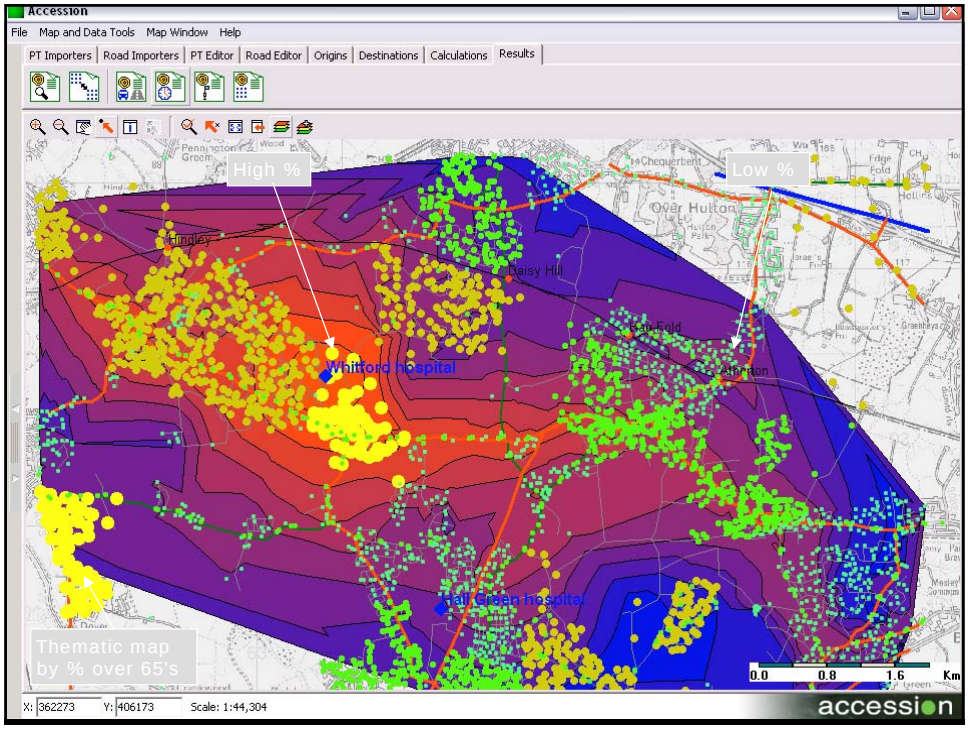
Isochrones

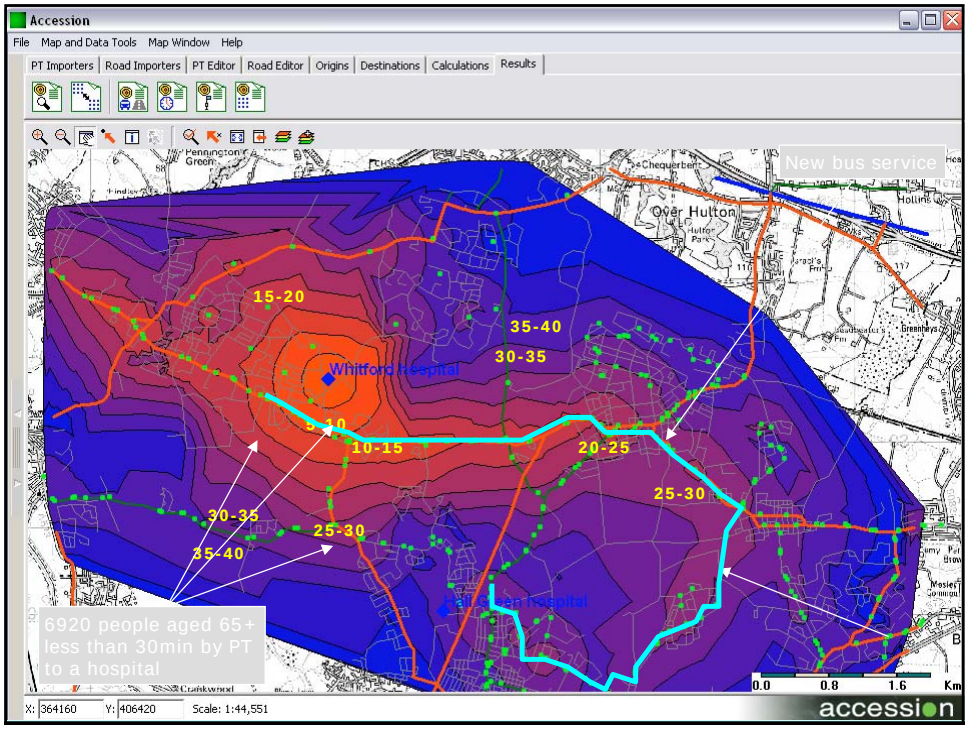
Journey times, in hours and including waiting times, from Edinburgh station at 7am on a weekday.




Accession
File Map and Data Tools Map Window Help
PT Importers Road Importers PT Editor Road Editor Origins Destinations Calculations Results
Whitford hospital
Hall Green hospital
20-25
35-40
30-35
5-10
10-15
15-20
25-30
8421 people aged 65+ less than 30min by PT to a hospital
0.0 0.8 1.6 Km
X: 364852 Y: 406478 Scale: 1:44,304
accession


Accession
File Map and Data Tools Map Window Help
PT Importers Road Importers PT Editor Road Editor Origins Destinations Calculations Results
Relocate services
Whitford hospital
20-25
35-40
30-35
5-10
10-15
15-20
25-30
20-25
30-35
35-40
5390 people aged 65+ less than 30min by PT to a hospital
0.0 0.8 1.6 Km
X: 364793 Y: 406384 Scale: 1:44,304
accession


Accession
File Map and Data Tools Map Window Help
PT Importers Road Importers PT Editor Road Editor Origins Destinations Calculations Results
High %
Low %
Whitford hospital
Hall Green hospital
Thematic map by % over 65's
0.0 0.8 1.6 Km
X: 362273 Y: 406173 Scale: 1:44,304
accession

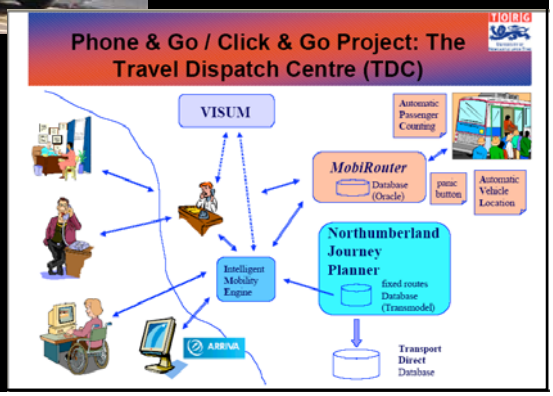

Accession
File Map and Data Tools Map Window Help
PT Importers Road Importers PT Editor Road Editor Origins Destinations Calculations Results
New bus service
15-20
35-40
30-35
Whitford hospital
10-15
20-25
25-30
30-35
25-30
35-40
Hall Green hospital
6920 people aged 65+ less than 30min by PT to a hospital
0.0 0.8 1.6 Km
X: 364160 Y: 406420 Scale: 1:44,551
accession

Demand responsive transport

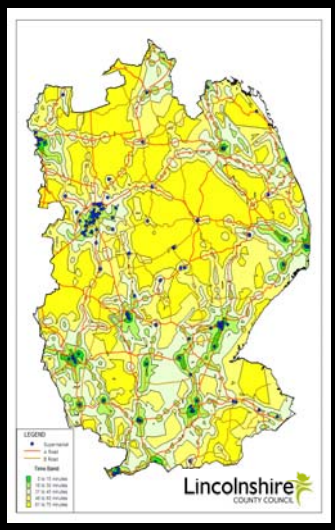
The Wigglybus has no set route. Its route is determined on a day-to-day basis by the passengers who use it.

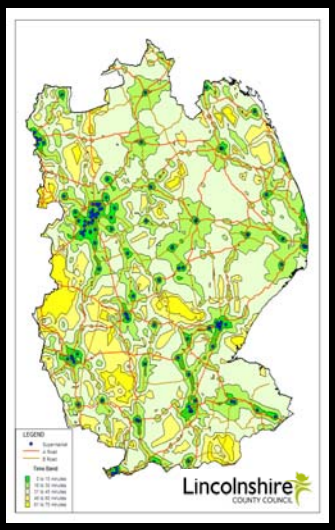
An experiment designed to tackle the problem of linking up a scattered population in an area with no obvious direct transport routes.

www.wigglybus.org.uk



Lincolnshire: impact of DRT

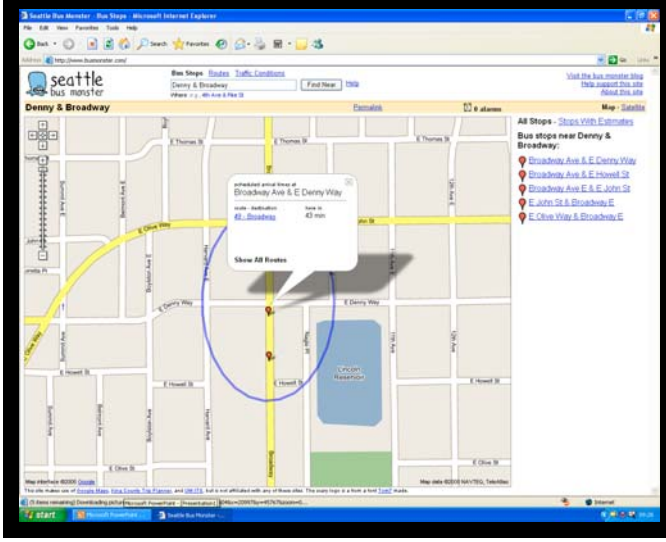




Access to supermarkets and service centres: before and after 'Call Connect'

Web mashing - (busmonster)

Alternatives to government as provider of “front-end” information?



Example: Bus Monster begins with Google Maps, overlays bus stops by scraping the King County Metro Trip Planner, adds real-time estimates of bus arrival times using the REST interface to University of Washington ITS, and adds traffic images using a number of Washington State web sites

PSA target: road safety

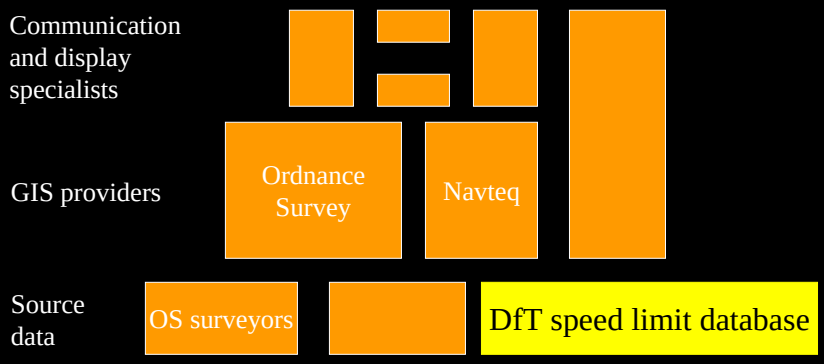
To reduce the number of people killed or seriously injured in Great Britain in road accidents by 40% … by 2010

Over a million GPS devices on the road (TRL estimate)

Major application: speed camera alert -

Speed limit database

Should the DfT make freely available a database of road speed limits in open format?

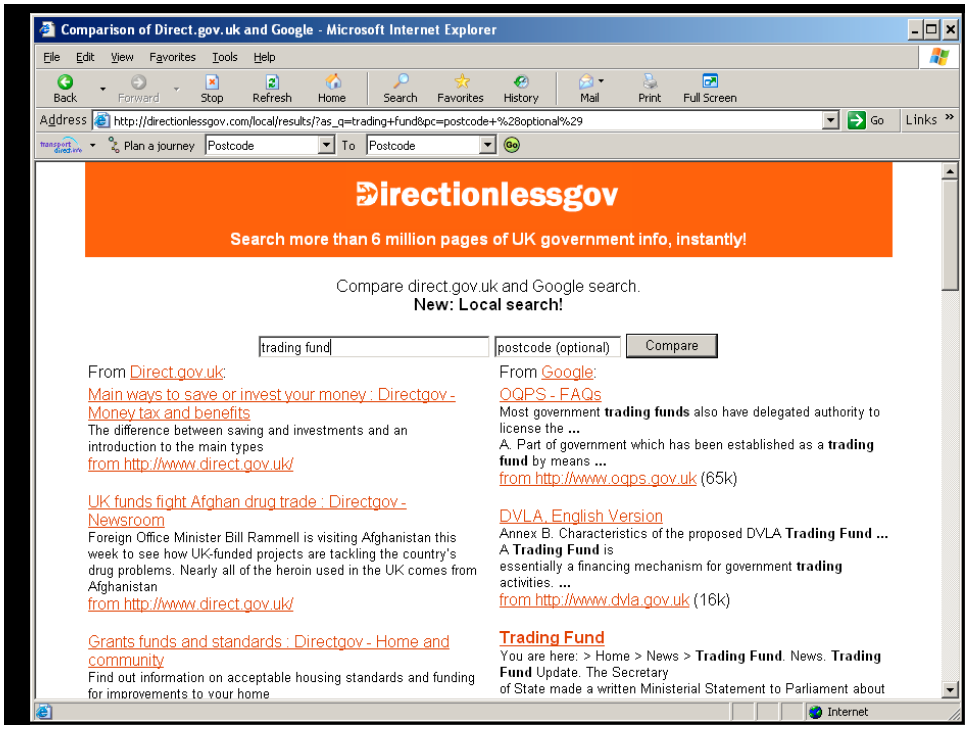


Value for money?

- Benefit-cost ratios of over 100
- OS has trading fund status
- Intellectual Property Rights

David Rhind: "There are some substantial potential benefits from having this and making it generally available, perhaps in cars: to make it useful, though, the coordinates of every section of road are needed and the obvious source is OS. Ideally, the DfT would like to make the entirety of this data – including the coordinates – freely available; that is, free from copyright and easily shareable, in the public domain. This appears to be enormously difficult."

www.foundation.org.uk/pdf18/fst18_9.pdf



New sources of transport data

www.payasyoudriveinsurance.co.uk

- smartcards, mobile phones, pay as you drive insurance
- trade-offs between privacy, convenience, personalisation

Google Earth: moving vehicles

Ikonos satellite: linear array CCDs, multiple images, taken at slightly different times and realigned on the ground. Thus double exposure for moving objects.

Challenge of real time data

- Growing amounts of real time and archived data across Government
- Improved IT systems allow "data mashing" – the "live" integration of data from diverse sources
- Offers potential of real time information and accountability
- Application to the **development**, **delivery** and **monitoring** of policies

Barriers to realising applications

- **Data sources** – what data is out there?
- **Data access** – how can data be more widely shared?
- **Data use** – what limits should be placed on data use and linking?
- **Data confidentiality** – how to protect / anonymize personal data?
- **Data security** – how to prevent unauthorised access?

The challenges

- **Science and Technology** – solutions and problems
- **Legal** – too much guidance, too little clarity?
- **Social** – attitudes depend on benefits?
- **Economic** – who pays and who benefits?
- **Political** – the conflicting roles of government

Science and Technology challenge

- *"When you can measure what you are talking about, and express it in numbers you know something about it; but when you can not express it in numbers your knowledge is of a meagre and unsatisfactory kind"* Lord Kelvin
- NSF report *"Revolutionizing Science and Technology through Cyberinfrastructure"* – ability to handle large amounts of data the most widely expressed need
- BUT for most government applications technical barriers are not the main obstacle - evolution in ICT creates barriers AND provides solutions to remove these barriers (e.g. data privacy, security…)

The Legal challenge

- Complex, poorly understood legislative regime
- Too much guidance
- The contractual relationship in data sharing
- Liability risk prevents innovation?
- The European dimension

Social challenge / trends

Increased monitoring / data collection

- Greater data exploitation by private sector (Tesco)
- Unpredictable data uses & applications

Changing public attitudes & trust (Foresight):

- Greater acceptance of personal data use?
- Acceptability depends on choice and the distribution of benefits?

Media perspectives

- **Your details are sold to car park extortionists**

The Government is at the centre of a 'secrets for sale' scandal after it admitted offering sensitive details about millions of motorists to private car parking companies.

Mail on Sunday, 20th November 2005 http://www.dailymail.co.uk/

- **Give us back our crown jewels**

Our taxes fund the collection of public data - yet we pay again to access it. Make the data freely available to stimulate innovation, argue Charles Arthur and Michael Cross.

The Guardian, March 9, 2006

http://www.freeourdata.org.uk

Economic challenge

How do we fund data collection and management?

- Data reuse and the information economy
- Trading funds – OS, Met Office, Land Registry…
- …but high data prices can prevent innovation
- Public money therefore public access?

How do we value data & information?

Direct charging vs indirect economic benefits

Political challenge

Government has several potentially conflicting roles:

- **Provider** – of public sector information
- **Custodian** – of public interest
- **Regulator** – of data use by public & private sectors
- **User** – for the development, delivery & monitoring of policies

Transformational Government Lab?

- Developing the algorithms and computer architectures necessary to handle and search large data sets is a major scientific challenge
- Government cannot replicate the capability of companies such as Microsoft and Google
- Government does have a role in enabling the delivery of innovative solutions by resolving issues that impede data access and use
- Experimentation with mashing (primarily of government data) in a public-private partnership could explore
  - the legal, social, economic and political issues
  - service delivery, efficiency

What will success look like?

- Better delivery of Government policy by improved accountability and decentralised decision making;
- Greater opportunities for public and private sectors to develop innovative solutions;
- Improved supply of information and services to the general public, how and when required.