

# Open Data Update for MRS CGG

November 2014

@oduguk



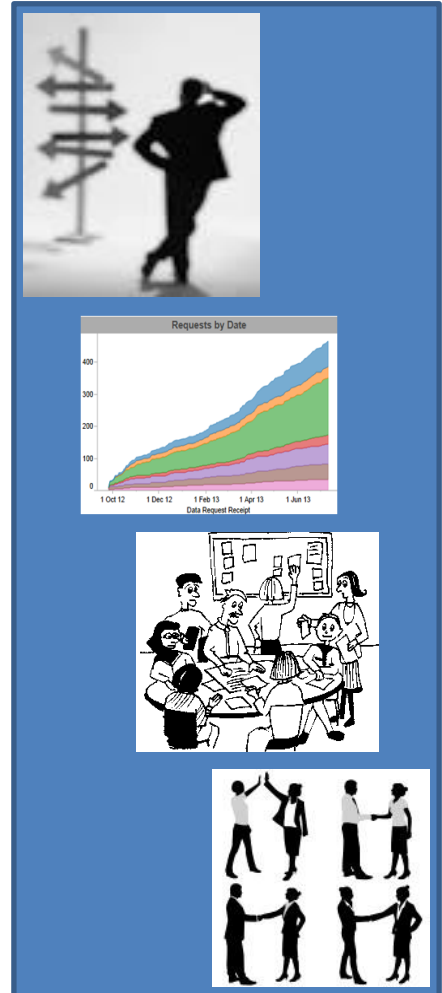
**DATA.GOV.UK**<sup>Beta</sup>  
Opening up Government



Cabinet Office

# ODUG remit

- **Outreach** - to Open Data users, re-users, wider stakeholders
- **Demand Led Open Data** - Prioritise user requests for data to identify the most important data sets for open release
- **Benefits Cases** - evidence gathering & research to inform cases for the release of Open Data
- **Release of Data Fund (RoDF)** – assess & prioritise bids and make recommendations for funding
- **Work with Central Government Sector Boards and the wider Public Sector**
- **Advise PSTB on Open Data policy and additional funding for Open Data**



# Open Data

- **Transparency**
  - holding government to account
- **Efficiency, innovation and enterprise**
  - Better public service delivery – *citizen information and choice*
  - Improved public sector efficiency – *joining the dots between spending and outcomes*
  - Economic growth - *opening up innovative UK market opportunities*



# Economic Value of Open Data

- **Estimated value of PSI to the UK economy**
  - £16bn per annum (UK Gvt, 2011)
  - More than £1.8bn of direct annual benefits p.a. *plus* wider social benefits of >£6bn (Shakespeare Review, 2013)
- **Global Revenues**
  - >\$3T annual global benefits (McKinsey, 2013)
  - Global revenues growing 13% p.a. (Oxera, 2013)
  - US Geospatial Industry \$73bn, 500,000 jobs. In US Drives \$1.6T revenue, \$1.4T cost savings (BCG, 2013)



# ODUG Achievements – Key datasets

- **Open Datasets released/committed**
  - HMLR Historic Price Paid and INSPIRE Polygon data
  - Register of UK Charities
  - EA Flood Data
- **Open Datasets on their way**
  - Historic Met Office Obs – *incremental release 2015*
  - VAT Register – *preliminary release & consultation*
  - DVLA Bulk & Police Stolen Vehicle Data – *consultation*
  - DECC NEED and EPCs – *work underway*
- **Datasets under debate**
  - Open National Address Dataset
  - Ordnance Survey Open Data



# ODUG Achievements - other

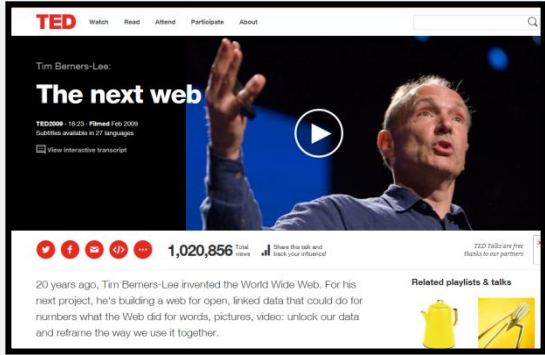
- **Demand Led Open Data – evidencing real-world demand**
- **Community Engagement – widely networked**
- **A voice for data users at the heart of government**
- **Active contribution to policy debate**
- **Release of Data Fund (RoDF) all funding now allocated**
- **National Information Infrastructure (NII) – ODUG review/analysis underway**



# National Information Infrastructure (NII)

- **The NII Intention:**
  - To ‘contain the data held by the government which is likely to have the broadest and most significant economic impact if made available and accessible outside of government’
  - October 2013; a published list of 233 datasets
- **ODUG reviewing what:**
  - *We should* know
  - The *value* of the NII might be
  - Data is *available* as either open or closed data
  - The *definition* of NII Core Reference Data

The economic potential of open NII data sources



20 years ago, Tim Berners-Lee invented the World Wide Web. For his next project, he's building a web for open, linked data that could do for numbers what the Web did for words, pictures, video: unlock our data and reframe the way we use it together.

*"Linked data could do for numbers what the Web did for words" (Berners-Lee)*



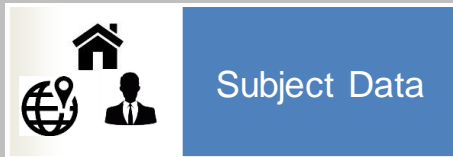
# Defining the NII

- Defining Infrastructure: *“The basic facilities, services and installations needed for the functioning of a community or society.”*
- Defining the Information Infrastructure: *“the core data assets describing this physical and digital infrastructure”*

## CORE NII DATA

The No1 priority at heart of the NII to unlock economic potential

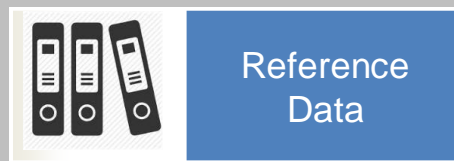
1. Data has unique identifiers necessary to reliably use, manage and monitor the subjects as well as enable connections to the other related data
2. Accuracy and quality must be maintained by the data controller to an appropriate level (ODI certificate standard level as a minimum)



The People, Organisations, Locations, Assets, Publications (eg legislation), Products and Services



Business Definitions, Technical Definitions and Quality Characteristics



Fundamental common reference data eg classification data, temporal and geographical data

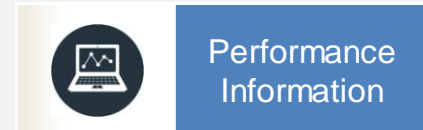
- Data should be open by default, and only closed where there are data privacy or security constraints
- Can be maintained by public organisations (eg Local Authorities) or private sector (eg Transport Operators)

## RELATED DATA TYPES

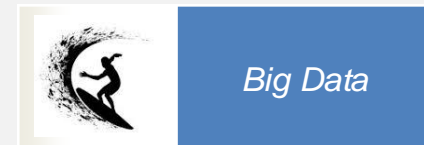
Potential for future NII data once core in place



Data captured about an event or transaction eg hospital visit, train journey, property sale. Captures new relationships between entities.



Aggregated information and reports illustrating KPIs eg operational MI, spending/ fiscal data or national statistics



In this context, raw data that can be used to improve the quality of a service eg social media content, sensor readings, web logs 8




# Health NII Illustration

NOTE: Open Data would not include personal identifiers, as Open Data would only be on an anonymised basis. This increases the importance of other non-personal identifiers e.g. location data to make links

## CORE NII DATA

## RELATED DATA TYPES



**Subject Data**

Premises/ Sites

- Hospitals
- GP Practices
- Health Centres
- Pharmacies
- Dental Practices
- Ambulance Stations
- Opticians
- Care Homes
- Pathology Labs

People

- Doctors (GMC)
- Nurses & Midwives (NMC)
- Therapists (HPC)
- Pharmacists (GPhC)
- Dentists (GDC)
- Opticians
- Paramedics
- Pathologists
- Trust executives

Products & Services

- Medical products (MH&A)
- In & Out patient procedures (OPCS)
- Mental Health Procedures (DSM4)
- Social Care Services (ESD)
- Grants
- Pathology reference data (NLMC)


Common reference data for linkages

What: Diagnoses types (ICD10), clinical terms (READ codes)

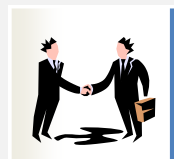
Where: Address data.

When: Time / period data

Who: *common Personal identifier*



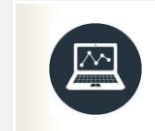
**Reference Data**



**Event / Transaction Data**

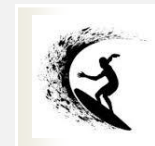
Hospital Episodes, Prescriptions, Care Visits, 999 calls

With core reference data, can link to other info eg weather/ transport disruptions



**Performance Information**

National Patient Survey(s), Aggregated Hospital Episode Statistics, Local Spending Data...



**Big Data**

Test results, Case notes, DNA sequencing

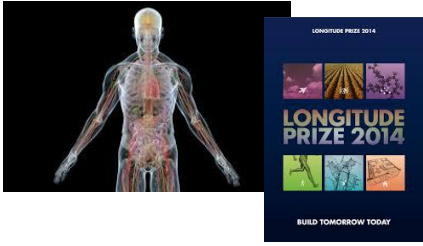
With core reference data can link to other info eg weather data

### Basic Gap Analysis: Current Health data in the NII

There are currently 32 datasets in the NII under 'Health & Social Care', currently 3 fit the above definition of Core NII Data - see Clinical Commissioning Group reference data (locations, size), The Casemix Service (healthcare activity classifications) and National Library Medicine Catalogue (pathology reference data)

# “Big Thinking”: What can the NII do?

## Leaps in Public Health



See Longitude Prize 2014 from NESTA  
Discovering connections and patterns from health and environmental data in order to tackle public health priorities (such as antibiotic resistance)

## Better Use of Public Services



**Avoidable Contact**  
Using data to identify preventative services to deflect demand away from expensive, reactive public services (eg A&E visits, law enforcement)

## Finding Economic Opportunities



**Better Investment Decisions**  
For any location: find available land, get business rates, available grants, nearest rail and motorway networks, median property values, employment levels, median incomes, speed of planning applications

## Social Enablement



**Quality of Life Indicators**  
Having national and local benchmarks to promote greater accountability and self-determination among communities

## Continuity Planning



**Mitigating Disruptions**  
What happened before when there was weather and/or transport disruption and how do health, transport, utilities providers, local authorities and communities cope better next time?

## Connected Public Services



**Better Service Design & Delivery**  
Better (more efficient and effective) service design and delivery helped by Tell Us Once-style information flows, reliable metrics and the resolution of gaps, overlaps and conflicts in data

# Core Reference Data - the “Where”, “What” and “When” to make connections

## Leaps in Public Health

See Longitude Prize 2014 from NESTA

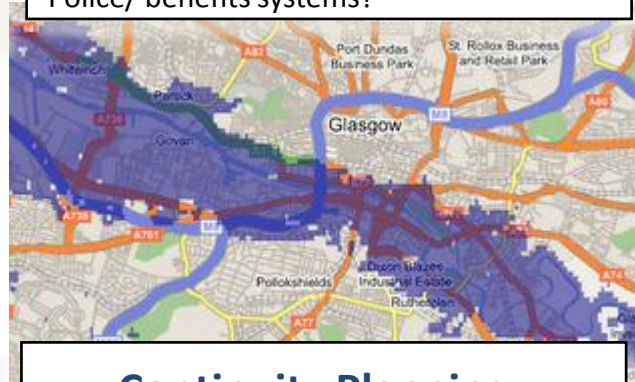
- Q. **What** has happened, **where** and **when**?
- Q. How does that compare to other areas?
- Q. What is distinctive about those **localities**?



## Better Use of Public Services

Avoidable Contact

- Q. **Where** are services located?
- Q. **Where** are the exemplars of Education, Health, Employment services around the country and why?
- Q. What is the demand on the **local** A&E/ Police/ benefits systems?



## Finding Economic Opportunities

Better Investment Decisions

- Q. **Where** is the available land for development?
- Q. What is the **proximity** to transport links, skilled workforces etc?
- Q. How successful have similar businesses, especially small local businesses been **here**?



## Social Enablement

Quality of Life Indicators

- Q. **Where** are my local public services?
- Q. How does their **proximity** and our population/ demographic compare to other communities?
- Q. What combinations work best and **where**? Can we replicate them?

## Continuity Planning

Mitigating Disruptions

- Q. **What** happened, **where** and **when**?  
E.g.: Which *roads/addresses* are prone to flooding? *Where* are accident blackspots?
- Q. **Where** are the local services and supplies?
- Q. **Where** should services be better located and configured?

## Connected Public Services

Better Customer Service

- Q. How can we be sure we are talking about the same **place/ location** between different organisations?
- Q. Who owns the single record of truth?

# Who needs the NII and why?



**Sarah**  
(and family)

## As a Citizen

- I need to know **how** and **where** I can get **access to Government services** (for Education, Health, Justice, Travel etc)
- I need aggregated information on operational **performance** and **financial spending** so I can hold elected officials **accountable**

## Connected Public Services

I want to be remembered

## Social Enablement

I want choice and participation



**Laura**

## As a Public Servant

- I need to understand and improve the design and quality of the services I provide
- I need to identify any cost or risk opportunities to the citizens I serve

## Better Use of Public Services

We want to deliver value for money

## Continuity Planning

We are learning the lessons



**Steve**

## As an Entrepreneur

- I need facts so I can identify a business opportunity to offer services **better, faster or cheaper**
- I want to offer services that are either social (about people and their interests), mobile (location based), or save people time

## Finding Economic Opportunities

I know where there is demand for my business



**Tony**

## As a Data Business Owner

- I need reliable, consistent sources of open, free data. Ideally via APIs with published SLAs. I can combine this data with customer, consumer or my own data to build a strong business proposition that delivers value to my users.

## Finding Economic Opportunities

I can create high value data products



**Rita**

## As a Researcher

- I need raw data that I can link so I can identify efficiencies across the organisations and services used by my customer groups

## Leaps in Public Health

I'm improving the lives of those with dementia

## Social Enablement

I want to innovate





# Ongoing challenges

- Ordnance Survey Open Data and Derived Data Licensing restrictions
- Agreeing the National Information Infrastructure (NII)
- More data releases; 'Open by Default'
- Getting public sector procurement right
- Rationalising the governance and availability of Open Data
- Establish UK position on data privacy
- Realising and measuring the full economic potential of Open Data
- Skills gap

