UN GWG GP Committee
Tech Partners

Working with Technical Partners

- Microsoft
- Amazon
- Demand Institute
- Esri
Global Platform: How we got here

UNSC Remit
World Data Forum
January 2017

Our Vision
A Global collaboration to harness the power of data for better lives [14/07/2017]

WORK STREAMS

- **WS1**
  Governance
  Define the business operating model for the GP and make the business case for the GP taking into account the legal structure of the GP, legal and regulatory compliance, overall governance, structure of partner agreements, the customer network and the options for long-term sustainable financing.

- **WS2**
  Data Policy and Common Technology Framework
  Draft and seek endorsement by the statistical commission of a data policy framework for governance and information management including ethics, privacy, confidentiality and security, which will shape the workings of the GP. Seek endorsement for a common technology infrastructure to enable its delivery.

- **WS3**
  Communications and Best Practice Sharing
  Seek endorsement by the Statistical Commission of a framework for communications and best practice sharing for the GP. Overall progress towards delivering the GP will be widely communicated to encourage new partners and users to join this initiative.
Joining up with other data initiatives
Statistical Modernisation @ONS
Data for ONS, Government and Researchers

Societal insight

Better Informed debate
Savings and efficiencies

Innovative economy

Economic insight

Better Informed research
Targeted Services
Targeted Service Delivery

Clustering

- Better policy decisions based on demographics and geography

Better Informed Public Debate: Migration

- Confused and mixed messaging from different sources
- Collaboration between ONS and OGDs
- Clearly present and explain all information

Service integration: DfE/CLG Land Availability Tool

- Identify possible sites for new free schools

Reduce reoffending rates

- Assess interventions

Better policy decision: Flow of funds

- Closer monitoring of financial flows
- Reduce risk of another financial crisis
- Asset and liability position by sector
- One sector’s liabilities are spread across economies
Data Infrastructure: delivering a critical resource

Government Information Infrastructure

- Data Catalogue
- Metadata
- Security / Access rights
- Analysis
- Open Data

Registers/Core Reference Data

- Local Government
- Central Government
- Census
- Citizens
- Business

Citizen-facing Services

Business-facing Services

Policy Design / Efficiency

Academic Research

Statistics/Economic Analysis

Innovation
The challenge: multiple data user lenses across estate

Statistical production:
Secure, in-house access
ONS staff

Statistical research:
Working in partnerships, project-based access

Statistical services:
Match, link, anonymise data: internal and 3rd party use

3rd party disclosure:
Accredited research and statistics
# National Statistician’s Data Ethics Advisory Committee (NSDEAc)

<table>
<thead>
<tr>
<th>ETHICAL PRINCIPLES</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Good</td>
<td>The use of data has clear benefits for users and serves the public good.</td>
</tr>
<tr>
<td>Confidentiality, data security, consent</td>
<td>The data subject’s identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is considered appropriately.</td>
</tr>
<tr>
<td>Methods and Quality</td>
<td>The risks and limits of new technologies are considered and there is sufficient human oversight so that methods employed are consistent with recognised standards of integrity and quality.</td>
</tr>
<tr>
<td>Legal Compliance</td>
<td>Data used and methods employed are consistent with legal requirements such as the DPA, the Human Rights Act, the SRSA and the common law duty of confidence.</td>
</tr>
<tr>
<td>Public views &amp; engagement</td>
<td>The views of the public are considered in light of the data used and the perceived benefits of the research.</td>
</tr>
<tr>
<td>Transparency</td>
<td>The access, use and sharing of data is <strong>transparent</strong>, and is communicated clearly and accessibly to the public.</td>
</tr>
</tbody>
</table>
Apprenticeship in Data Analytics

- 2 year programme
- Level 4 Diploma-Data Analytics

MSc Data Analytics: Government

- September 2017
- Multiple academic partners
- Framework reviewed by GSS

Government Data Accelerator

- Open to Public sector staff
- Run with GDS & GO-Science
- Network “hub” @ DSC

Research Teams

- Launched September 2016
- 24 FTE
- 60 FTE projected March 2018

Partnerships

MoUs agreed including Universities (Cardiff), research institutes (Alan Turing), international statistical institutes (Stats Netherlands)
Collaboration with national & devolved government, e.g. DEFRA, DCMS, DFID, Welsh Government
ONS and National Addressing

Local Authorities
Royal Mail
Ordnance Survey

Common secure services available to public sector
(and potentially wider subject to licensing)
common service > consistency & efficiency

GeoPlace → UPRN

Addressing Service

Match
Address
Validate
Batch
Validated Address

HS
ONS Address Matching Service

ONS
Citizen Service

Feedback to source
(Improving quality)

Matching for the public sector

ONS Data Management Platform

GeoPlace
AddressBase

Business Index
Address Index
Business Index

**Input**

- **Current Data**
  - Companies House
  - VAT
  - PAYE

- **Future Data**
  - Self assessment
  - Corporation Tax
  - Charities and Commissions

**Function**

- **Business Index**
  - Matching & Linking algorithm: Combines records into linked data (legal units)

**Output**

- **Business Index**
  - Legal unit spine

**Company details**

- Companies House
- VAT
- PAYE

**Data Science**
Data Integration – 3 way index

01
Does not currently exist so requires linkage to be carried on a case-by-case basis

02
ONS delivers the Inter-Departmental Business Register (IDBR) across government. To be referenced to the Business Index

03
ONS maintains a register of addresses and is working with the Government Digital Service for use across government
IDEAS – A New Data Model

1. Business Index Index
2. Address Index
3. Statistical Index
The Material Properties of a dataset

<table>
<thead>
<tr>
<th>CONTENT</th>
<th>What the data describes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SENSITIVITY</strong></td>
<td>Secret</td>
</tr>
<tr>
<td><strong>IDENTIFIES</strong></td>
<td>Individuals</td>
</tr>
<tr>
<td><strong>GRANULARITY</strong></td>
<td>Field Variable</td>
</tr>
<tr>
<td><strong>RECENCY</strong></td>
<td>Real-Time</td>
</tr>
<tr>
<td><strong>RELIABILITY</strong></td>
<td>Complete</td>
</tr>
<tr>
<td><strong>RELEASE</strong></td>
<td>Closed</td>
</tr>
<tr>
<td><strong>AUDIENCE</strong></td>
<td>Named</td>
</tr>
</tbody>
</table>
WHY ARE WE HERE TODAY?

User needs

Infrastructure

Collaboration

Data Access