Tech Partners

Working with Technical Partners

Microsoft

Demand Institute

Amazon

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

Better Informed research

Targeted Services

Economic insight
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

Efficiency

• The case for change
  - 500,000 cases are sorted each year for those who face the law
  - 21 new community resettlement services
  • Innovation and efficiency to make a difference
  • Paid by results
  • Investment in services to prevent reoffending

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

Service integration: DfE/CLG Land Availability Tool

• Identify possible sites for new free schools
Data Infrastructure: delivering a critical resource

Local Government
Central Government
Census
Citizens
Business

Government Information Infrastructure

Registers/Core Reference Data

Data Catalogue
Metadata
Security / Access rights
Analysis
Open Data

Citizen-facing Services
Policy Design/Efficiency
Academic Research
Innovation

Business-facing Services
Statistics/Economic Analysis
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
Data use and access

**DATA IN**
- **STATUTORY**
  - SRSA (inc ISOs)
  - DEA
  - STA (business)
  - RSA (registration)
  - VAT/finance acts
- **NON-STATUTORY**
  - Voluntary Surveys
  - Non-controlled admin data
  - Commercial partnerships
  - Open data

**PREPARE**
- **METADATA**

**ANALYSE**
- **STATISTICAL METHODS DEVELOPMENT**
  - 3rd Party Service
  - 3rd Party Disclosure
  - Devolved Statistics

**INFORM**
- **ACCREDITED RESEARCH OUTPUT**
  - Inform Policy
  - Ad Hoc Outputs
  - Statistical Releases

**DATA OUT**
Data Capability at ONS

National Statistician
John Pulinger

Deputy National Statistician for Data Capability
Heather Savory

Digital Services & Technology
David Best
- Digital Publishing
- Digital Technology (IT Infrastructure/Products, Platforms & Computing)
- Enterprise Architecture & Service Design
- Digital Policy & Service Standards
- Information Assurance & Technical Security
- Systems & Data Security
- Operations Support & Maintenance
- Service Delivery & Design
- Technological Policy & Standards
- Technical Service Policy

Business Services & Development
Neil Wooding
- Corporate Planning and Resilience
- Human Resources
- Learning Academy
- People and Physical Security
- People Capability
- Project Design and Delivery

Data Science Campus
Tom Smith
- Cross-Gov Data Science
- Data Science Frameworks & Definitions
- Data Science Policy & Standards
- Data Science Technical Delivery & Design
- External Partnering
- GSS/Heads of Profession Leadership

Methods, Data & Research
Sarah Henry
- Data Policy and Standards
- Information Infrastructure
- Methodological Policy & Standards
- Methodological Services (GSS)
- Research Accreditation / ADRN & VML
- Statistical Quality Centre
- Data Services
- Good Practice and GSS support
## National Statistician’s Data Ethics Advisory Committee (NSDEC)

<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>
Active Learning

- **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

Experimentation

- **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
ONS Address Matching Service

Feedback to source (improving quality)

Matching for the public sector

ONS Address Matching Service

SINGLE
ADDRESS -> API
UPRN ←

 BATCH
ADDRESS -> API
UPRN ←

GeoPlace
AddressBase

ONS Data Management Platform

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 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

**CONTENT**

<table>
<thead>
<tr>
<th><strong>SENSITIVITY</strong></th>
<th>Secret</th>
<th>Private Personal</th>
<th>Commercial</th>
<th>Open</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IDENTIFIES</strong></td>
<td>Individuals</td>
<td>Groups</td>
<td>Non Personal</td>
<td></td>
</tr>
<tr>
<td><strong>GRANULARITY</strong></td>
<td>Field Variable</td>
<td>Record</td>
<td>Sample</td>
<td>Population</td>
</tr>
<tr>
<td><strong>RECENTY</strong></td>
<td>Real-Time</td>
<td>Periodic</td>
<td>Historical</td>
<td></td>
</tr>
<tr>
<td><strong>RELIABILITY</strong></td>
<td>Complete</td>
<td>Substantial</td>
<td>Patchy</td>
<td>Incomplete</td>
</tr>
<tr>
<td><strong>RELEASE</strong></td>
<td>Closed</td>
<td>Restricted</td>
<td>Open</td>
<td></td>
</tr>
<tr>
<td><strong>AUDIENCE</strong></td>
<td>Named</td>
<td>Closed Group</td>
<td>Third Parties by Type</td>
<td>Public</td>
</tr>
</tbody>
</table>
WHY ARE WE HERE TODAY?
Thank you
Heather Savory
Deputy National Statistician & Director General for Data Capability
heather.savory@ons.gov.uk
@SaturnSA4