

CIVIL SERVICE QUARTERLY

Issue 13
Jan 2017

FEATURE
PUBLIC SERVICES AND
THE NEW AGE OF DATA

Subscribe for free here:
quarterly.blog.gov.uk
#CSQuarterly

GROWING A GOVERNMENT
DATA SCIENCE COMMUNITY



Civil Service

CONTENTS

PUBLIC SERVICES AND THE NEW AGE OF DATA	John Manzoni Cabinet Office Permanent Secretary and Chief Executive of the Civil Service	4
GROWING A GOVERNMENT DATA SCIENCE COMMUNITY	Paul Maltby - Former Director of Data, Government Digital Service, Sue Bateman - Deputy Director, Better Use of Data, Cabinet Office	8
BETTER STATISTICS, BETTER DECISIONS	John Pullinger UK National Statistician, Head of the Government Statistical Service (GSS) and Chief Executive of the UK Statistics Authority	12
USING DATA AND DESIGN TO HELP PEOPLE STAY IN WORK	Catherine Drew Former Senior Policy Designer, Policy Lab, Cabinet Office	16
HEALTHY INFORMATION SHARING	Imogen Fuller Engagement Manager, Centre of Excellence for Information Sharing	20
OPEN DATA FOR BETTER OUTCOMES	John Seglias Chief Technology Officer, Department for Environment, Food and Rural Affairs	22
BUILDING AN OPEN SOLAR POWER MAP	Michel Blancard Data Scientist, Etalab, French Government	24
DATA ANALYTICS FOR MORE EFFICIENT SERVICES AND BETTER LIVES	Andrew Goodman Programme Director, Home Office Digital, Data & Technology Capabilities & Resources	28
BUILDING AN INTERNATIONAL COALITION TO COMBAT ANTIMICROBIAL RESISTANCE	Professor Dame Sally Davies - Chief Medical Officer for England, Nicole Redhead - Global Health Policy Manager - International AMR, and Tracy Parker - UK AMR Strategy Programme Coordinator	30
DATA IN THE LIFE OF...	Charlie Boundy Head of Data Science, Department for Work & Pensions Aimee Murphy Data Scientist, Home Office	32
TRANSPARENCY AND EVIDENCE - SHOW YOUR WORKINGS	Dr David Halpern Chief Executive of the Behavioural Insights Team (BIT)	35

Civil Service Quarterly opens up the Civil Service to greater collaboration and challenge, showcases excellence and invites discussion. If the Civil Service is to be truly world-leading, it needs to collaborate more, learn from experts outside the Civil Service, listen more to the public and front-line staff and respond to new challenges with innovation and boldness.

Any civil servant can write for Civil Service Quarterly - contact csq@cabinetoffice.gov.uk

Cover: DESIGN102

DESIGN102
The GCS Design Centre

Contact us at:
design102@justice.gsi.gov.uk
www.design102.co.uk

CONTACT US

csq@cabinetoffice.gov.uk
Room 140, 70 Whitehall,
London, SW1A 2AS

Read the magazine online
and subscribe for free -
quarterly.blog.gov.uk

EDITORIAL TEAM

Adam Thorndike, Cabinet Office
adam.thorndike@cabinetoffice.gov.uk

Simon Holder, Cabinet Office
simon.holder@cabinetoffice.gov.uk

Aneesa Kazi, Cabinet Office
aneesa.kazi@cabinetoffice.gov.uk

Bethan Godwin, Cabinet Office
bethan.godwin@cabinetoffice.gov.uk

Thanks to Claire Pringle,
Department of Health

DESIGN BY DESIGN102

Sanita Bajwa, Creative Designer
Darren Poole, Creative Designer

EDITORIAL BOARD

Chris Wormald, Permanent Secretary,
Department of Health (chair)

Chris Skidmore, Minister for the Constitution

Alex Aiken, Executive Director,
Government Communications

David Halpern, Chief Executive,
Behavioural Insights Team

Clare Moriarty, Permanent Secretary,
Department for Environment, Food
and Rural Affairs

Sir Richard Lambert, Chairman,
Board of Trustees, British Museum

Paul Maltby, Director of Data,
Government Digital Service

John Pullinger, National Statistician
and Chief Executive of the UK Statistics
Authority

Dave Ramsden, Director General,
Chief Economic Adviser, HM Treasury

Charles Roxburgh, Director General,
Financial Services, HM Treasury

Jill Rutter, Programme Director,
Institute for Government

EDITORIAL

Civil Service Quarterly: Data in government



Chris Wormald

Welcome to the 13th edition of Civil Service Quarterly.

The theme of this edition is data in government and the opportunities - and challenges - it presents for creating smarter policy and transforming the efficiency and effectiveness of public services.

We feature three articles from senior Civil Service figures:

- **Public services and the new age of data** - John Manzoni, Chief Executive of the Civil Service and Cabinet Office Permanent Secretary
- **Growing a government data science community** - Paul Maltby, former Director of Government Data in the Government Digital Service (GDS), and Sue Bateman, Deputy Director, Better Use of Data, Cabinet Office
- **Better statistics, better decisions** - John Pullinger, UK National Statistician, Head of the Government Statistical Service (GSS) and Chief Executive of the UK Statistics Authority

Four articles describe how different departments and teams are applying data to provide better public services or releasing it for productive use outside government:

- **Using data and design to support people to stay in work** - Catherine Drew, former Senior Policy Designer, Policy Lab, Cabinet Office
- **Healthy information sharing** - Imogen Fuller, Engagement Manager, Centre of Excellence for Information Sharing
- **Open data for better outcomes** - John Seglias, Chief Technology Officer, Department for Environment, Food & Rural Affairs
- **Data analytics for more efficient services and better lives** - Andrew Goodman, Programme Director, Home Office Digital, Data & Technology Capabilities and Resources

A data scientist at the French Government equivalent of GDS contributes the first in what we hope will be a series of articles showcasing case studies of innovation and excellence in overseas civil services:

- **Building an open solar power map** - Michel Blancard, Data Scientist, Etalab, French Government

Moving away from the data theme for a moment, we focus on the UK's leading role in tackling a critical issue for global health:

- **Building an international coalition to combat antimicrobial resistance** - Dame Sally Davies, Chief Medical Officer for England, Nicole Redhead, Global Health Policy Manager,

International AMR, and Tracy Parker, UK AMR Strategy Programme Coordinator

We have also spoken to two data specialists about their experience of being a data scientist in government in:

- **Data in the life of...** - Charlie Boundy, Head of Data Science, Department for Work & Pensions; and Aimee Murphy, Data Scientist, Home Office

Finally, we feature a personal view of the arguments for opening up to wider scrutiny the data and analysis behind government decision-making:

- **Transparency and evidence - show your workings** - Dr David Halpern, Chief Executive, Behavioural Insights Team

I hope you enjoy this issue. You can give us your views and comments on the Civil Service Quarterly blog (<https://quarterly.blog.gov.uk/>), by email (csq@cabinetoffice.gov.uk), or via #CSQuarterly on Twitter. If you would like to submit an idea for a feature in a future edition, please get in touch.



Chris Wormald
Permanent Secretary,
Department of Health

PUBLIC SERVICES AND THE NEW AGE OF DATA

John Manzoni, Permanent Secretary at the Cabinet Office and Chief Executive of the Civil Service, describes how government is changing to exploit the vast potential of data for making better policy and better services.



John Manzoni

Government holds massive amounts of data. The potential in that data for transforming the way government makes policy and delivers public services is equally huge. So, getting data right is the next phase of public service reform. And, as can be seen in this edition of Civil Service Quarterly, the UK Government has a strong foundation on which to build this future.

Public services have a long and proud relationship with data. In 1858, more than 50 years before the creation of the Cabinet Office, Florence Nightingale produced her famous 'Diagram of the causes of mortality in the army in the east' during the Crimean War. The modern era of statistics in government was born at the height of the Second World War with the creation of the Central Statistical Office in 1941.

HOW DATA CAN HELP

However, the huge advances we've seen in technology mean there are significant new opportunities to use data to improve public

services. It can help us:

- understand what works and what doesn't, through data science techniques, so we can make better decisions: improving the way government works and saving money;
- change the way that citizens interact with government through new better digital services built on reliable data;
- boost the UK economy by opening and sharing better quality data, in a secure and sensitive way, to stimulate new data-based businesses;
- demonstrate a trustworthy approach to data, so citizens know more about the information held about them and how and why it's being used.

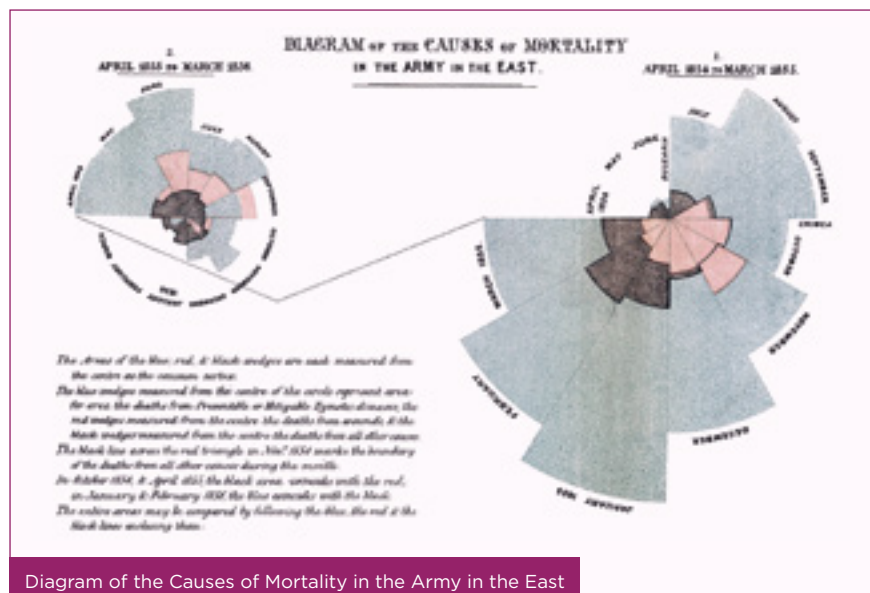
In 2011 the Government embarked upon a radical improvement in its digital capability with the creation of the Government Digital Service (GDS), and over

the last few years we have seen a similar revolution begin on data. Although there is much more to do, in areas like open data, the UK is already seen as world-leading.

'BIG DATA', TOOLS AND TECHNIQUES

This new digital age of data is utterly changing how we live our lives, how we shop, socialise and consume media. At the heart of this are a range of new tools, techniques and types of data, often rather misleadingly bundled together as 'Big Data'.

The scale of some data in government can be extremely large (weather data, for example). However, in practice, much of the innovation in public services and data will come through applying new machine learning tools and techniques to the existing data that can be queried and transported around our systems in



more modern ways.

The reality is that the whole fabric of government is changing as it becomes digital. More and more data is capable of flowing around government and between government and users. There's now the potential to access vast amounts of data and powerful tools to help us analyse and use it.

NECESSARY CHANGES

But if government is going to seize this opportunity, it needs to make some changes in:

- **infrastructure** - data is too often hard to find, to access and to work with; so government is introducing developer-friendly open registers of trusted core data, such as countries and local authorities, and better tools to find and access personal data where appropriate through application programming interfaces (APIs) for transformative digital services;
- **approach** - we need the right

policies in place to enable us to get the most out of data for citizens and ensure we're acting appropriately; and the introduction of new legislation on data access will ensure government is doing the right thing – for example, through the data science code of ethics;

- **data science skills** - those working in government need the skills to be confident with data; that means recruiting more data scientists, developing data science skills across government, and using those skills on transformative projects.

POWERING DECISIONS IN THE FRONT LINE

One thing to highlight is how the status of data, and of those working with it, is changing in government. Analysis has too often been seen as the preserve of a policy elite; something for ministers and senior boards rather than the life-blood of operational decision-making in

government. And while there has been a steady growth in the use of business intelligence data across operational parts of government, with these new data science tools and techniques we are entering an age when analysis will increasingly be built into new digital services; powering decisions made in the moment by frontline workers. This is a substantial shift, and one that will lead not only to greater efficiency, but also to a more personalised experience of government for citizens.

Data is the foundation of government, a part of our essential national infrastructure, and it cannot be left to chance. The data revolution has shaken entire industries such as retail, transport and financial services, and this disruption is coming for government too.

John Manzoni

Permanent Secretary at the Cabinet Office and Chief Executive of the Civil Service

DATA SCIENCE - WHAT IT CAN DO FOR GOVERNMENT

DATA TOOLS AND TECHNIQUES

The government has an ambition to deliver 500 new free schools within this Parliament, and appropriate sites are needed on which to build these schools.

Data scientists at Government Digital Service combined data on land from a range of sources to create a tool that identifies possible sites for new schools. The tool uses a bespoke algorithm to rank the sites based on a number of variables, such as number of pupils, proximity to green space and transport. It provides a simple way of checking what land is available, based on specific criteria. This will save time and money when trying to identify building sites, potentially speeding up the creation of new schools.

The screenshot shows the 'Land Availability Finder' tool on the GOV.UK website. At the top, it says 'GOV.UK Land Availability Finder'. Below that, a 'DISCOVERY' banner reads: 'This is a new service – your feedback will help us to improve it.' The main heading is 'Find available sites'. There are three main sections for user input:

- What do you want to build?**: Two radio button options: 'Primary School' (selected) and 'Secondary School'.
- How many pupils?**: A text input field.
- In what location?**: A text input field with a link 'use current location'.

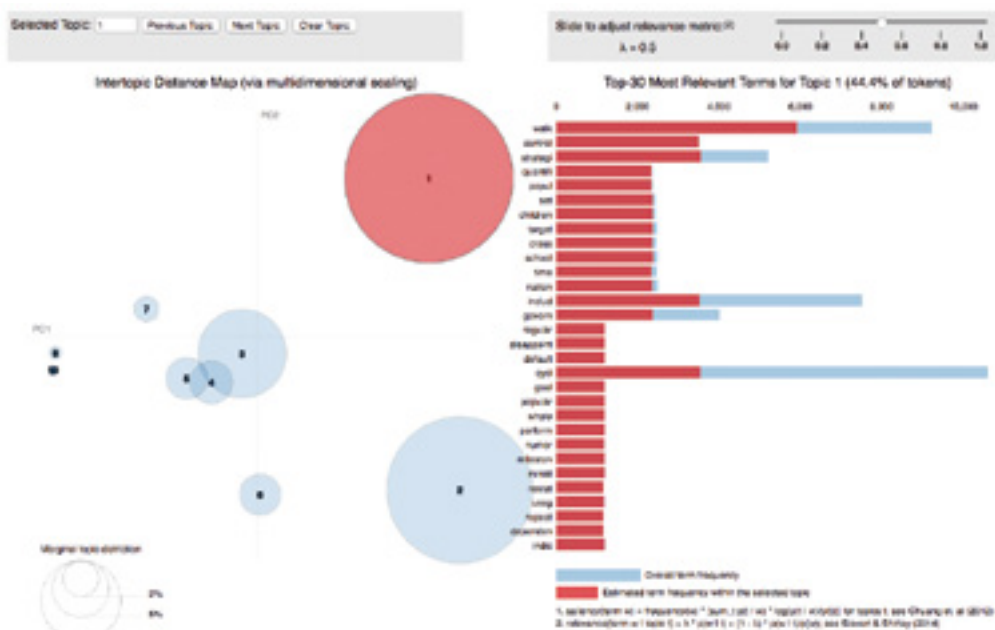
At the bottom, there is a map showing a geographical area with various locations marked.

MACHINE LEARNING

Government holds consultations when it is thinking of implementing a particular policy and asks the public for views. Much of the collation and analysis of feedback is currently done manually, which is time-consuming and prone to error.

The Department for Transport (DfT) is developing a tool that automates the collation process. The tool will

extract relevant data from emails and export it into a data structure. The tool uses automated machine learning and natural language processing to examine and process the content of the responses into an interactive topic model, which allows policy colleagues to explore the data.



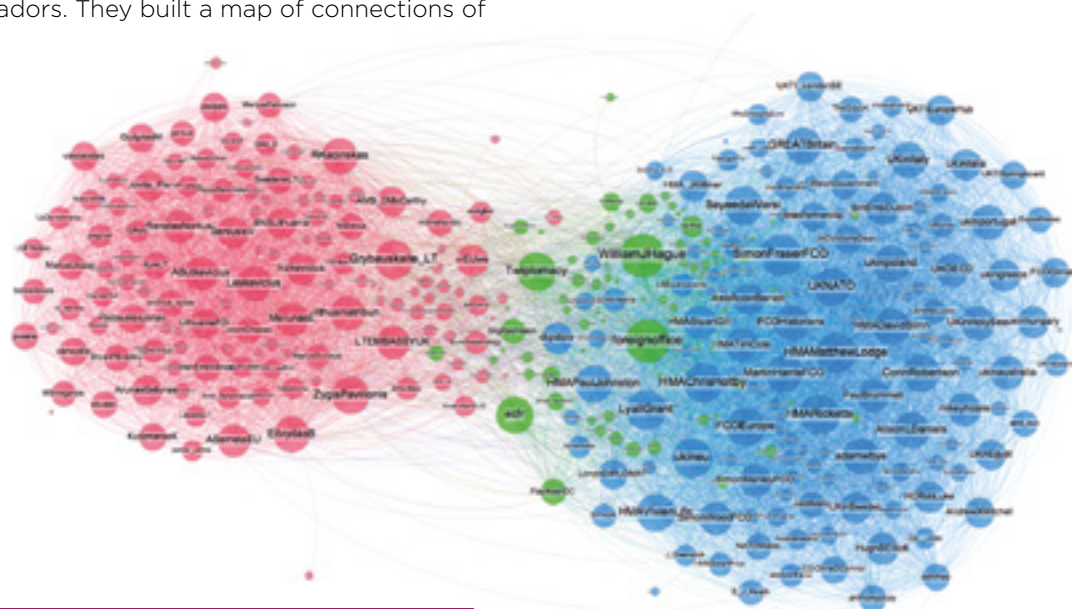
Screenshot of DfT data collation tool

DATA VISUALISATION

The Foreign & Commonwealth Office (FCO) wanted a clear picture of their online presence and to have sight of key influencers and important conversations on social media.

Using open source tools, data scientists created visualisations of the social media networks of British ambassadors. They built a map of connections of

Twitter followers and sorted them into groups representing their affiliations. The visualisation gives insight into each ambassador or embassy's influence online, as well as the locations and interests of their followers. This insight helps the FCO tailor its content to be relevant to followers.



@HMAvilnius (Ambassador to Lithuania) influence on Twitter

TOP 10 DATA SCIENCE TERMS - AND WHAT THEY MEAN

Algorithm

At its simplest, an algorithm is a step-by-step list of one or more mathematical operations applied to some data and giving an output – for example, the sum of two numbers.

Big data

Although difficult to define precisely, ‘big data’ is commonly thought of as data that is so large, complex or rapid that it cannot be processed or analysed on a single computer alone.

Data mining

This describes the process of deriving insight from data by using computers. It is now largely superseded by the term ‘data science’, referring to a set of skills and tools for exploring data to – for example – identify trends in behaviour or predict the likelihood of various outcomes.

Deep learning

A type of neural network (see *below*) that is capable of identifying and learning

patterns from very large amounts of data. For this reason, it is widely used in image recognition systems.

Machine learning

Describes algorithms that are able to learn patterns from data, and use these to make predictions when presented with new data. Historically, machine learning is divided into two branches: ‘supervised’ and ‘unsupervised’ learning (see *below*).

(Artificial) Neural network

A supervised machine-learning algorithm modelled on the structure of the brain. Despite being developed in the early 20th century, neural networks only gained prominence in recent years as computing resources have become cheaper and more widely available.

Statistics

A branch of mathematics concerned with describing or predicting patterns in data in the presence of uncertainty. The boundaries between statistics and machine learning are

somewhat blurred. Certain algorithms are claimed by both fields, whilst others (for instance, neural networks) fall more squarely under machine learning.

Supervised learning

Machine-learning algorithms that must be told in advance what patterns to look for. These algorithms can be extremely powerful for a wide range of tasks including handwriting, speech recognition, and image classification.

Unsupervised learning

Machine-learning algorithms that can look for patterns in data, without advanced knowledge of what patterns may exist in that data. ‘Clustering algorithms’ are an example of unsupervised learning algorithms.

Web scraping

Automatically extracting data from websites by developing software to make multiple queries of a website or web service much more quickly than a human would be able to do.



GROWING A GOVERNMENT DATA SCIENCE COMMUNITY

Paul Maltby, former Director of Data, GDS, and Sue Bateman, Deputy Director, Better Use of Data, Cabinet Office, chart the development of data science capability in government.

The first small group to bring data science skills into government was formed with a clear imperative: to build a community of practical knowledge within government so that it had sufficient skills to begin a process of transformational change.

There was also a recognition that, with this bedrock of skills, government could engage with private data science providers from a position of strength.

From 2011, there had been a revolution in digital and technology in UK central government. This saw more agile digital expertise come back in house, bringing an end to a period where skills and strategy were outsourced to a small number of systems integrators.

During this period, government's role was often relegated to little more than contract management, mopping up after too many expensive and protracted failures.

To capture the best of the new data agenda and fit it within digital services – without losing the ability to know what and when to buy from the market – government had to develop its own community of data science expertise. But there were two big challenges: the first was that the number of known data scientists in government could be counted on the fingers of one hand, and recruitment of these sought-after individuals would be difficult and expensive; the second was a lack of awareness – and even some defensiveness – among existing analytical professions.

This is the story of how this joint project team overcame these hurdles, developed a community in government of more than 350 individuals with a data science capability, and started to put this capability to use to drive value for citizens.

KICKING THINGS OFF

The UK was an early world leader in open data. It released non-personal data sets collected by government in machine-readable formats, for no cost and with licence terms that permitted anyone to use (or even sell) the data as they saw fit. This activity has driven real-world applications – from how we find information about the next train or bus, to how businesses manage due diligence.





Those who witnessed this rapid change at first hand were struck by the different tools and approaches external data companies and civic society activists brought to government data, compared to those used by the government's own sizeable (and excellent) community of statisticians, economists, social and operational researchers. They knew that in our personal lives, data-driven digital companies had already utterly transformed how we shop and socialise.

“Data science work soon became a critical component of the cross-government data programme led from GDS. That programme grew out of an understanding that to maximise the reform potential of this agenda it would not be enough merely to increase data science capacity.”

The Cabinet Office and Government Office (GO) for Science horizon-scanning team had also picked up on this technological shift and government's lack of capability in the area. With the Cabinet Secretary's support they set up a small project team to explore the potential of this new approach to data within government. The Cabinet Office Innovation Group, the GDS, the GO for Science combined under the leadership of the Economic and Domestic Secretariat (EDS) Director General and with guidance from the operational research and government statistical professions to get things rolling.

The initial strategy had four parts, which have stood the test of time over the subsequent three years:

- to 'show not tell', by doing some practical demonstration projects, as opposed to writing strategy papers to explain in the abstract what the project might mean;
- to find and grow data science capability, and to broaden the understanding of what value these new tools can provide;
- to overcome practical and technical barriers; for example, common data science tools such as the programming languages 'R' and 'Python', not being accessible from many government computer systems;
- to ground this work in an ethical approach that, from

the start, aimed to consider what we should do with these potentially powerful tools, not just what we could do with them.

Data science work soon became a critical component of the cross-government data programme led from GDS. That programme grew out of an understanding that to maximise the reform potential of this agenda it would not be enough merely to increase data science capacity. Instead, this capability had to be grounded in the new digital services being built across government.

In addition, government's often woeful data infrastructure had to be fixed. One way was with trustworthy sources of core reference data, such as countries, local authorities and schools, through open registers; and the construction of a scalable system for appropriate personal data exchange through APIs (application programming interfaces), which allow data sets across departmental boundaries to be queried, rather than shared in bulk. This, in turn, would require an updated policy and legislative framework, not only to remove unnecessary friction (through data access provisions in the Digital Economy Bill, for example), but also to put in place new rules and procedures, for instance, on the ethical application of these new tools.



Break-out group at GDS

BUILDING A COMMUNITY

On the capability side, the original strategy was a combination of hiring small amounts of relatively junior data science talent from outside government, and complementing this by developing digital skills and culture within the existing government analytical community. The aim was always to build local hubs of expertise across government, rather than trying to form a single central team in GDS or elsewhere.

This community would be built around the nucleus of the small handful of existing data scientists within GDS and the Office for National Statistics (ONS), and others with the necessary skills elsewhere in government. Finding them was easily done in some cases. For example, with a few calls to colleagues in the intelligence communities, or connecting with the world-leading expertise in GCHQ and the Defence Science and Technology Laboratory (DSTL). However, pockets of expertise existed in unexpected areas. There were some truly impressive individuals and work going on in places such as the

Health and Safety Executive labs in Buxton, Derbyshire, and Bootle, on Merseyside.

These early enquiries revealed a degree of untapped interest. However, given the aim of bringing analysis out of the shadows and putting it centre-stage in government, there was also a surprisingly cautious response in parts of the established analytical community. While many clearly relished the opportunity to update their skills, there were some who dismissed the new data science agenda as “trying to pretend it invented maths” and claimed data science had been practised in government since the time of the experimental physicist Patrick Blackett and the amazing innovations in operational research during and since World War II.

The first group of around 30 data scientists was assembled in February 2014. They met to discuss the emerging data science programme and provide ideas for demonstration projects to show the potential of data science to a wider network of senior leaders in government.

A plethora of ideas and experiments circulated in this

early phase, with some noble and necessary failures (Cabinet Office ran some lunchtime coding clubs for both analysts and non-analysts that proved quite popular but not especially useful). However, in building the now thriving community of data scientists, three types of activity proved particularly telling.

First, spreading the word about our data science ambitions by presenting to as many public service organisations and professional boards as possible, and blogging about early prototypes and intentions. This included a number of senior seminars in the main departments, with Permanent Secretaries and their analytical, digital/technology and policy leaders, chaired by Sir Mark Walport, the Government Chief Scientific Adviser. A surprise in some of these sessions was that the senior analysts were unaware of the lack of connections between their digital and technology counterparts and their work.

Use of wider communications channels and presentations, such as sessions at Civil Service Live, conferences of the different professions, including policy

as well as the formal analytical professions, and at numerous external events was also effective in spreading the word of data.

With the data in government blog, the aim was to foster connections between existing talent and to focus the attention of senior leadership on the data agenda. Many departments produced data science action plans in these early days. These plans were a useful lever to concentrate minds on the projects that might best add value, bring together various communities within departments, and provide space and permission for experimentation, though this would rarely produce significant value at the start.

The second area that helped create the data science community was the Data Science Accelerator, a programme for ambitious government analysts wanting to grow their budding data science skills. The accelerator linked them up with data science mentors across government. In the early days, it also provided them with a working Macbook Pro, free of the sometimes overly enthusiastic government IT restrictions that put everyday data science tools out of their reach. In return, the government analysts brought a practical project and sufficient time to devote to their learning, including spending a day together as a community every week in GDS's Holborn office and, as the programme developed, in three hubs around the country in London, Sheffield and Newport.

The Data Science Accelerator was designed as a minimum

viable training programme, which was spun up quickly and iterated rapidly, rather than as more strategic, longer-term training for existing analysts. Lying squarely in the remit of the existing professional communities, this more routine training is now also really beginning to show results.

The third activity that made a difference in developing the government data science community was the efforts made to connect this community in practical ways, from the bottom up. That meant regular 'meet-ups', gatherings where data scientists across the public sector could show early progress in their projects and exchange ideas and practice. They were also opportunities for guest speakers from outside government to inspire and show how these same tools were being used in the private sector.

From these early beginnings to build a community has grown a wide range of opportunities to share and connect. These include a dedicated messaging app, where code and frustrations can be shared, and an assortment of data drinks, lunches and dedicated community groups within departments.

Along the way, data scientists kept popping up. They were not always formally described as such, and their unique talents were not always being put to the best use - asking data scientists to grind through regular statistical releases from a department is a too common waste of their time and talent. However, the number of data scientists in government continues

to grow as a result of hiring, training and uncovering previously unknown talent. There are now more than 350 identified individuals with these skills in the public sector.

THE FUTURE...

The development of this community and the work they are producing is a great Civil Service success story. It shows how quickly government can adapt, even without the relentless market pressures of the private sector.

There are now tens of data science case studies - a mixture of prototypes (some successful, some less so) and increasingly serious value-adding propositions. The UK Government is one of the best in this field, and significantly ahead of many larger organisations in the corporate sector, though it still lags behind some of the groundbreaking digital-first companies.

The government data science community has emerged from its early development phase. It stands ready to build its insights into serious transformational digital services, and to use these new techniques to complement the traditional analytical advice to boards, programmes and ministers in increasingly sophisticated and visual ways. It can be increasingly confident of its ability to partner with the UK's amazing private sector data science companies where useful.

Paul Maltby - Former Director of Data, GDS, **Sue Bateman** - Deputy Director, Better Use of Data, Cabinet Office



BETTER STATISTICS, BETTER DECISIONS

John Pullinger, UK National Statistician, Head of the Government Statistical Service (GSS) and Chief Executive of the UK Statistics Authority, describes the role of GSS in mobilising the power of data to help government make better decisions that improve people's lives.

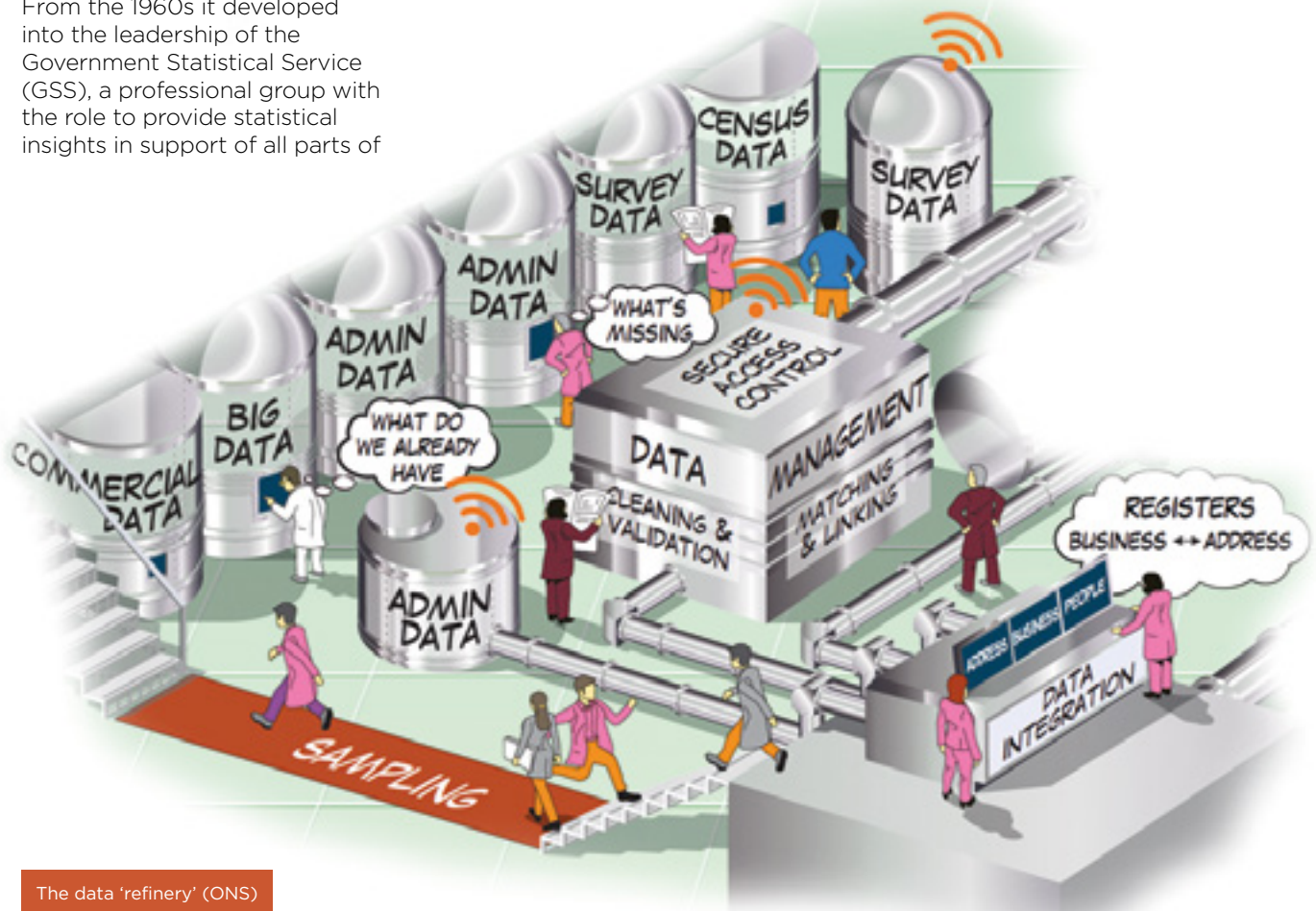
Official statistics have long been at the heart of decision-making in the UK. In 1941, Winston Churchill created the post of National Statistician at the centre of government. Its function was to create a body of information that would be accepted and used without question by all parties to a decision.

The initial focus was on the war effort. After the war, the emphasis was on macro-economic measurement and the development of a well-articulated system of national accounts. From the 1960s it developed into the leadership of the Government Statistical Service (GSS), a professional group with the role to provide statistical insights in support of all parts of

government. During this period, social statistics grew to maturity alongside economic data.

Into the 1990s, the open government agenda, initially championed by John Major when Prime Minister, saw official statistics as an essential element in democracy, serving government and the wider community. The 2007 Statistics and Registration Service Act enshrined this broad function to serve the public good in statute.

The official statistics system in the UK therefore has a proud tradition, reflecting the root of the word statistics in the idea of the state - the information needed for a state or a community to function well. But the world has moved on again, driven by the digital world we live in and the new sources of data available to us. The challenge now for official statistics is how to adapt to this new reality.



The data 'refinery' (ONS)

THE STRATEGY FOR STATISTICS

The response has been to adopt a strategy that calls for radical change. The role of the GSS is to mobilise the power of data to help make better decisions that improve people's lives. Those involved in decision-making – governments, businesses and citizens – have ever higher expectations. They need data faster, in more fine-grained forms, targeted at the questions that need answers. The statistical community must organise to meet – or even exceed – those expectations.

We have moved from a

position of data scarcity to one of data abundance. Statisticians must seek out and use all sources of data to help generate the insights that decision-makers are seeking. There are many people offering data and statistical analysis, and the GSS needs to distinguish itself as offering quality, trustworthiness and value. Its role is to make sure the figures are understood, accepted and used without question in a rapidly changing environment. To do so, it is necessary to develop the demand side (helping policy makers and the public be better equipped to use data) and

the supply side (presenting figures in ways that are easy to interpret, through visualisation, for example).

PRIORITIES

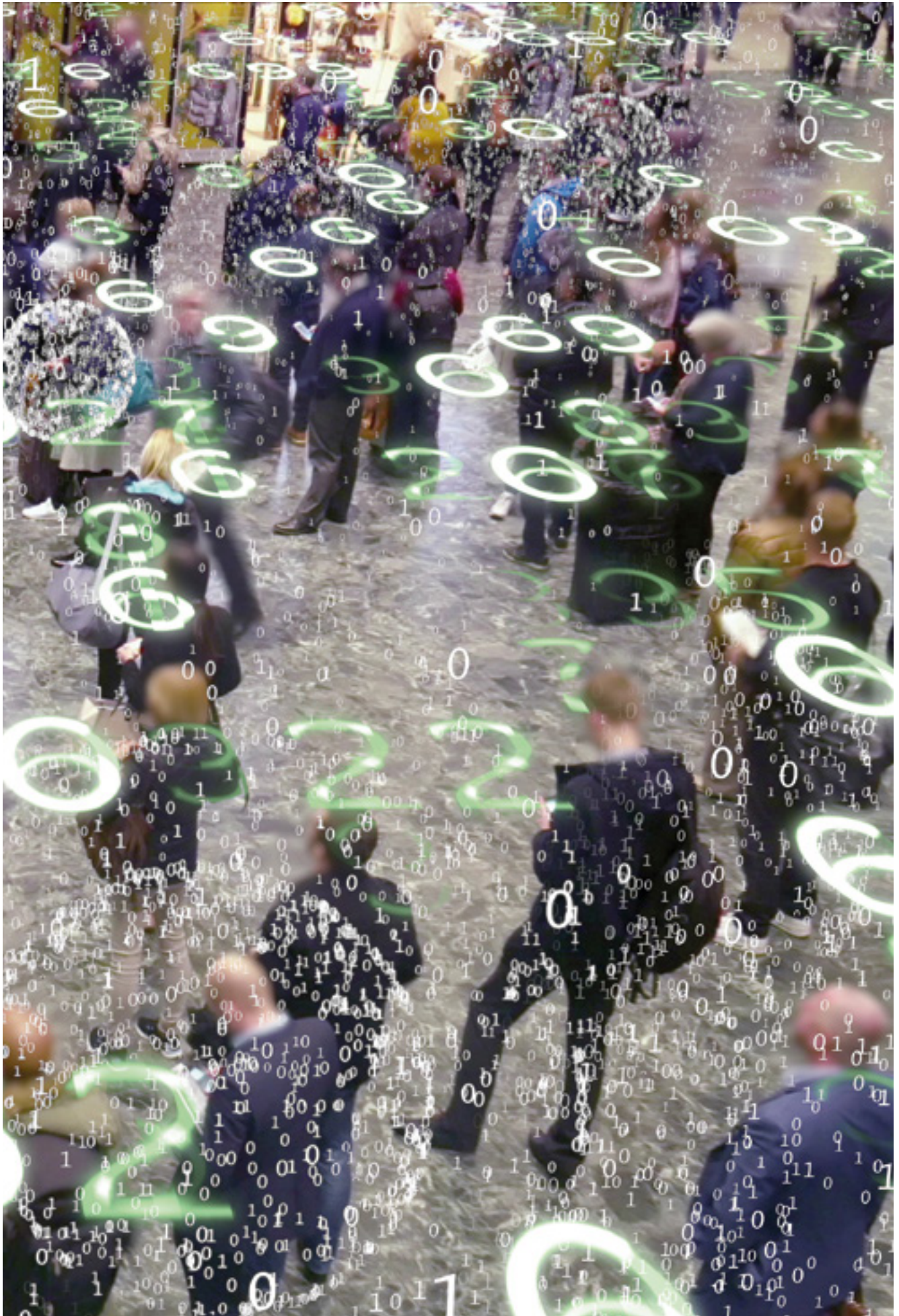
The delivery of this strategy needs to reflect the needs of the country. The top priority has been measurement of the economy. Guided by an in-depth review led by Sir Charles Bean, there is a comprehensive work programme to generate better understanding of the modern economy, including through a new Economic Statistics Centre of Excellence.

ECONOMIC STATISTICS CENTRE OF EXCELLENCE

The Economic Statistics Centre of Excellence will provide a source of expertise for the analysis of emerging and future issues needed to better measure the modern economy. The centre will offer the capacity to develop new methods and statistics. This will help us meet the challenge of measuring new forms of economic activity in a globalised world, develop new local, city and regional statistics and help us understand the UK's productivity performance. The new centre will also support discussion and dialogue around official statistics.

On 14 November, the Office for National Statistics (ONS) announced the consortium of leading economic institutions that will form the Economic Statistics Centre of Excellence. A total of eight institutions will come together to create the centre. This includes King's College London, the LSE (London School of Economics and Political Science), innovation foundation Nesta, NIESR (National Institute of Economic and Social Research), University of Cambridge, Stanford University in the USA, Strathclyde Business School, and Warwick Business School. The centre was scheduled to start work in January 2017.





The second priority has been to support public policy, joining up the efforts of statisticians and other analysts across organisational boundaries. There are strengthening levels of engagement between statisticians and both policy and operational colleagues. Alongside this there is a systematic effort to draw in and work with partners outside government who can provide support in serving the public good. Information is being brought together from various sources to develop a better-rounded understanding of complex issues. Examples of this include education, justice and migration. The flagship project for the GSS is a new approach to the Census, which holds the prospect of rich data for small areas being available much closer to real time.

INVESTING IN CAPABILITY

Underpinning all this is investment in data capability across government. This means both much better technology and the development of new skills and techniques. The opening this year of a Data Science Campus (DSC) at the ONS headquarters in Newport is a symbol of this endeavour and heralds substantial benefits to come.

Two years into the strategy for statistics, there is confidence it is making progress but a feeling that we have to raise our ambition higher still. The imperative in the post-EU-referendum environment is even greater (and more urgent) than in 2014. Statisticians (and the wider analysis and insight function) in government have always had a key role to support better decisions across government. Businesses and other organisations that have grasped the potential of data are prospering. The Civil Service is well placed to do so, too.

John Pullinger

UK National Statistician,
Head of the Government
Statistical Service (GSS) and
Chief Executive of the UK
Statistics Authority



ADMINISTRATIVE DATA CENSUS

ONS's Administrative Data Census aims to replicate the range of information collected through the Census using administrative data already held by government, supplemented by surveys. With access to administrative data sets, and if this new approach proves feasible, it would allow census-style outputs at lower cost and more frequently.

ONS is publishing annual research outputs and assessments of progress, and plans to benchmark the progress made on the administrative data census against the results from the 2021 Census. This will allow users to assess the quality and completeness of the alternative and will inform decisions about the future of the Census.



DATA SCIENCE CAMPUS

The DSC opened its doors in October 2016 in temporary space in the ONS offices in Newport, South Wales. Thirty staff will be in place by the end of 2016/17. In March 2017 the DSC will relocate to its permanent home on site, with capacity for up to 60 staff.

An Apprenticeship in Data Analytics has been launched, with the first group of apprentices starting work at the end of 2016. This two-year vocational training programme is the first of its kind in the UK. The DSC will also support the development of an MSc in Analysis for Government, with a dedicated data science pathway, for a September 2017 launch.

The main role of the DSC will be to conduct short data science projects with impact. The current team of seven data scientists has already begun work on the first two projects, with results expected early in 2017.

USING DATA AND DESIGN TO SUPPORT PEOPLE TO STAY IN WORK

Catherine Drew, a former Senior Policy Designer at Cabinet Office’s Policy Lab, describes how the lab has applied new data methods to help people manage health conditions and remain in work.

There are around 2.5 million people on health-related benefits in the UK, which costs some £15 billion per year¹. Each month, around three times as many people move onto health-related benefits as move off; and the longer people are on these benefits, the less likely it is that they will return to work.

The Work & Health Unit (a joint unit set up by the Department for Work & Pensions (DWP) and the Department of Health) has been working with Policy Lab for the past 18 months on a project to create ways to support people to manage their health conditions and stay in work.

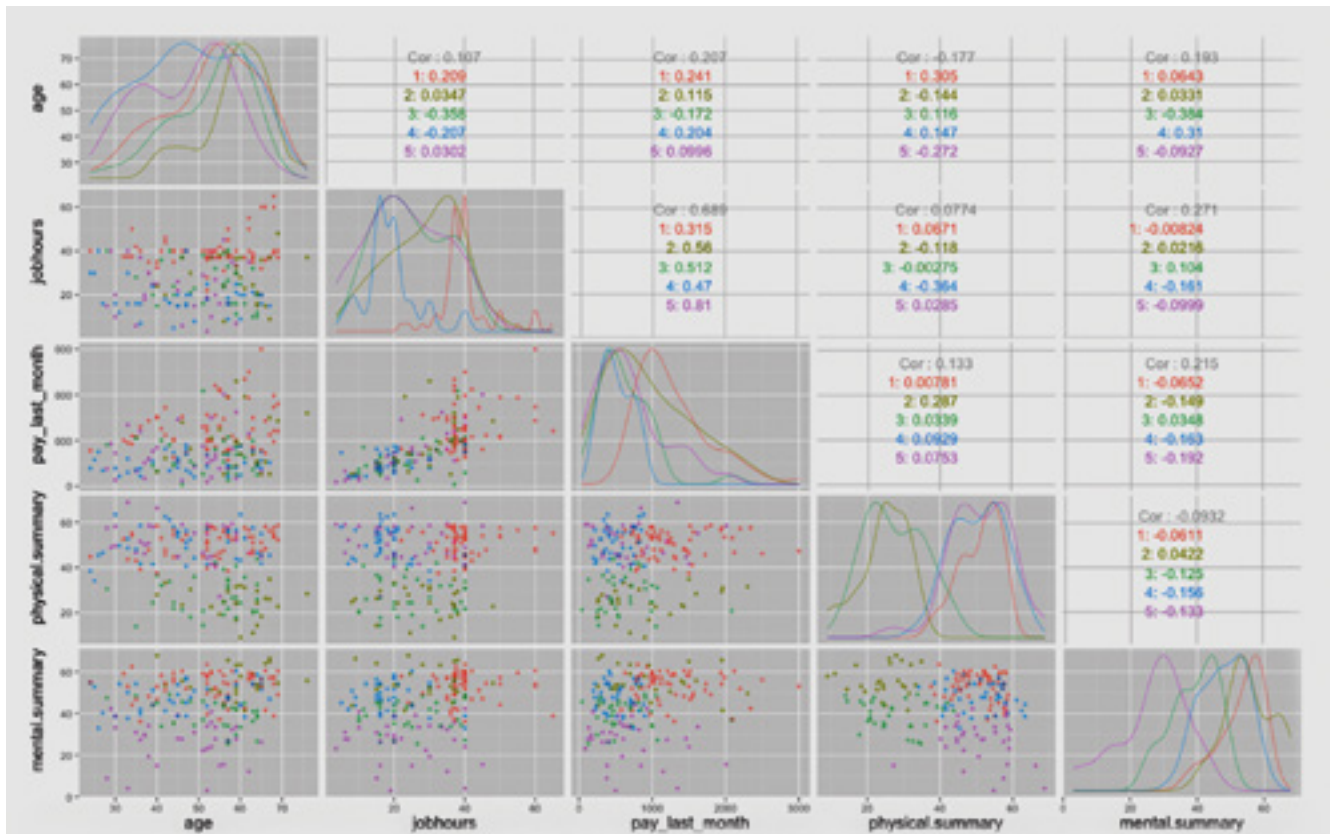
POLICY LAB AND THE THREE ‘D’S

Policy Lab is an innovation team in the Cabinet Office and supports departments to use new digital, design and data (DDD) techniques to improve how they make policy. It was created as part of the open policy-making agenda. Over the last two years it has not only supported departments on real policy projects that have saved money and improved citizen experience, but also used these as test-beds to create better policy-making processes and trained up to 5,000 civil servants

in new methods.

Data and digital are fairly understandable concepts in policy-making. But design? Why is it one of the three Ds?

Policy Lab believes that design approaches are particularly suited to complex issues that have multiple causes and for which there is no one, simple answer. Design encourages people to think about the user’s needs (not just the organisation’s needs), brings in different perspectives to innovate new ideas, and then prototypes (mocks them up and tries them out) to iteratively improve ideas until they find one that can be scaled up.



Segmentation analysis of those who reported being on health-related benefits in the Understanding Society survey



Policy Lab also recognises that data alone cannot solve policy problems, and has been experimenting with how to combine numerical and more human practices. Data can explain what is happening, while design research methods – such as ethnography, observing people’s behaviours – can explain why things are happening. Data can be used to automate and tailor public services; while design means frontline delivery staff and citizens will actually know about and use them. Data-rich evidence is highly valued by policy-makers; and design can make it understandable and accessible to a wider group of people, opening up policy-making in the process.

The Lab is also experimenting with new data methods.

Data science can be used to look at complex, unstructured data (social media data, for example), in real time. Digital data, such as social media data or internet searches, can reveal how people behave (rather than how they say they behave). It can also look at huge amounts of data far quicker than humans, and find unexpected patterns hidden in the data. Powerful computers can identify trends from historical data and use these to predict what might happen in the future.

SUPPORTING PEOPLE IN WORK PROJECT

The project took a DDD approach to generating insight and then creating ideas. The team (including the data science organisation Mastodon C and design agency Uscreates) used data science techniques together with ethnography to create a rich picture about what was happening. Then it used design methods to create ideas for digital services with the user in mind, and these were prototyped and tested with users.

The data science confirmed many of the known risk factors, but also revealed some new insights. It told us what was happening at scale, and the ethnography explained why.

- The data science showed that people were more likely to go onto sickness benefits if they had been in the job a shorter time. The

ethnography explained that the relationship with the line manager and a sense of loyalty were key factors in whether someone stayed in work or went onto benefits.

- The data science showed that women with clinical depression were less likely to go onto sickness benefits than men with the same condition. The ethnography revealed how this played out in real life. For example, Ella [not her real name], a teacher from London, who had been battling with depression at work for a long time but felt unable to go to her boss about it. She said she was “relieved” when she got cancer, because she could talk to her boss about a physical condition and got time off to deal with both illnesses.
- The data science also allowed the segmentation of groups of people who said they were on health-related benefits. Firstly, the clustering revealed that two groups had average health ratings, indicating that other non-health-related issues might be driving this. Secondly, it showed that these two groups were very different (one older group of men with previously high pay and working hours; the other of much younger men with previously low pay and working hours). The conclusion was that their motivations and needs to stay in work - and policy interventions - would be different.
- The ethnography highlighted other issues that were not captured in the data but would be important in designing solutions, such as: a lack of shared information across the system; the need of the general practitioner (GP) to refer patients to other non-health services as well as providing a fit note; and the importance of coaching, confidence-building and planning.

BUILDING A SERVICE

The Policy Lab team ran a co-design session with users, GPs, employers, Jobcentre Plus staff and policy-makers to turn this rich insight into ideas. Solutions formed around a Health and Work coach who could signpost

people to different non-health services, liaise with their employer to make adjustments and build their confidence.

The team knew it could not build a new work and health service from scratch. So, local areas prototyped elements of it in real-life situations (‘experience prototyping’) to see how it could fit with existing services. The elements tested included: a way for employers to refer their employees to the service (Penzance); for GPs to refer their patients to it (Southend); the role of the coach (East London); and a Health & Work book for users to keep their information all in one place (Bournemouth).

“Something like this would have been useful while I was still in work. I wasn’t as quick as some other staff because of my condition but they didn’t understand that. It might have helped me talk to my manager better about my health and what help I needed from him.”

[Response from a user in Penzance, testing the employer referral element.]

The next stage was a four-day hackathon with DWP digital, taking the feedback from testing to build a digital version of the service. Policy thinking had developed by this point, as the Work & Health Unit had also been drafting the Health, Work & Disability Green Paper. The agile nature of the project (that is, characterised by short phases of work and frequent reassessment and adaptation of plans as necessary) meant that we were able to use this prototype to explore and test new priorities: a three-way conversation between employer, coach and employee, and increasing self-referrals



Online Health & Work Plan

to Health & Work coaches via digital services. Bringing policy-makers, coders, UX (user-experience) specialists, frontline workers and users together meant that we were able to create a digital service with trust and empathy between employers and employees designed in.

WHAT WAS LEARNED?

There are a few key learnings from this project that the Policy Lab team has either applied to other projects (on homelessness, for example) or acted on as a result.

Firstly, there is huge power in using data science and ethnography together, with the data science giving the broad, large-scale 'what' and the ethnography providing the rich 'why'.

Creating multi-disciplinary teams from the start is essential to making sure that the

emerging hypotheses from the ethnographers and data scientists are shared. It means they can easily investigate further, and policy-makers can add their own expert interpretations.

Secondly, prototyping is a really important way of thinking through how a policy will be delivered from the outset. Creating a service blueprint (a plan of how someone will become aware of, join and use a service), 'experience prototyping' and testing a digital version gets policy-makers to think about how their ideas will actually work in practice when delivered by the frontline and experienced by citizens.

Finally, there are a few practical issues to consider, including giving external contractors access to government data, and getting non-data expert policy-makers to understand the potential for data science. The data science teams in Government Digital Service

are leading a programme of work to increase data science skills, capability among the analyst community, and data literacy amongst policy-makers so they can ask for it.

Policy Lab itself has developed a range of workshops, tools and techniques that can bring analysts and non-data-literate policy-makers together to help understand the potential of data, where it is generated from and how it could be used.

Read more about Policy Lab and its work at: <https://openpolicy.blog.gov.uk/>

Catherine Drew

Former Senior Policy Designer, Policy Lab, Cabinet Office

1. Black, C. and Frost, D. (2013) Health at Work: an independent review of sickness absence; Annual Report of the Chief Medical Officer

HEALTHY INFORMATION SHARING

Imogen Fuller, Engagement Manager, Centre of Excellence for Information Sharing, describes how local communities around England are changing the way they share health and social care information, and how this approach is helping to improve services to patients.

Across the country, health and social care providers are facing pressure to deliver improved, integrated services to more people, with fewer resources. The challenge for these organisations is to make the most of what they have and focus their efforts where they can deliver the maximum positive impact for the people they care for.

This is where the Centre of Excellence for Information Sharing (the Centre) comes in: working with providers in England to unlock one of their biggest assets - information.

Sharing information has the potential to make a real difference to services and the people who access them, but it is paramount to ensure it is used appropriately and that any confidential information is properly safeguarded. Sue Bateman, Deputy Director of Better Use of Data in GDS, and chair of the Steering Group for the Centre, said: "Effective information sharing can lead to improving services at lower cost, but some of the perceived barriers mean

that this can still be challenging. The work of the Centre is crucial in helping overcome some of those barriers, leading to a greater understanding of the potential of information sharing and an increase in collaborative working."

A key challenge for local initiatives is to ensure that patients are supportive of the necessary changes in the information sharing that supports joined-up work. Their permission to use patient data is critical, but there is a balance to be struck with the potential to create society-wide benefits from this data. As well as being informed enough to feel comfortable in consenting to share their information, patients should be able to see the tangible benefits it delivers.

THE VALUE OF GOOD INFORMATION

Why is information such an asset? The more effectively that providers can share information with partners, the more detailed the picture they can build of patients and the local population, allowing them to tailor their support and target their services at those who need them most.

One approach to maximise this untapped data is to introduce a 'common electronic record', in the form of an Integrated Digital Care Record (IDCR). In simple terms, this is a single point of reference for some or all of a patient's data. This can then be accessed and, in some cases, added to by the different health professionals involved in that person's care.

This move from a siloed approach to a single shared record was initially seen by local areas as a way to improve services and save money. However, additional benefits have been noticed, such

as allowing partnerships to reflect on and evaluate the culture in which they work, inform the way they can deal with future service demands and fully realise the opportunities that IDCRs offer.

Early adopters of IDCRs have already tackled the main cultural barriers to information sharing - such as leadership, communication, relationships and trust - helping them to work even more closely with partners.

As with all activity that involves sharing personal information, the feelings of the individuals whose data is being shared has to be central. The challenge is to implement the necessary changes to provide an improved service and better outcomes while reassuring the public.

CONVERSATIONS ARE KEY

Leeds is one area to have put information sharing at the heart of integrating healthcare services, highlighting early and ongoing citizen engagement as critical to its success.

The commitment to citizen input led to the formation of 'Joined Up Leeds'. This enabled conversations between local people and health and social care providers, giving an insight into people's motives when deciding whether or not to share information. This helped to understand local viewpoints on information sharing both in health and social care and the wider context.

The conversations began with influential local thinkers to provide a baseline of current experience of information sharing. They were then widened to a larger audience of local people, through carefully chosen channels, to improve accessibility.



This engagement identified public opinions that needed to be addressed. These included the perception that information might be sold to private companies (raised by 71% of participants); or that it might not be stored securely (a concern for 61% of participants). Busting these myths helped to create a groundswell of positivity for information sharing.

Leeds has now created a solid evidence-base of public opinion – the Leeds Care record – to call on when designing their IDCR, and set some important early communication messages for local people to help alleviate public scepticism.

PUTTING PUBLIC OPINION AT THE HEART OF SERVICE REFORM

Health and social care for the 2.1 million residents of North West London is provided by a Whole Systems Integration Care (WSIC) programme made up of 30 health and social care organisations.

From the outset, WSIC put patients and local community engagement at its heart. In particular, they wanted to ensure that service integration met people's needs through lay partners (the term used locally for patient, carer and community representatives), who are engaged in two ways.

They did this, firstly, through discussions at higher board level to define the overall framework for integrating care; and then through a wider forum.

This gave a voice to patients and carers, drawing particular attention to what is working and what could be improved, ensuring decisions about the services offered and the way information is shared are influenced directly by the end-users.

CONVERTING PUBLIC NEGATIVITY

Southend is another area where information sharing is a key component in creating integrated health and social care with a single, comprehensive care package for a patient's needs.

When a vocal opponent of the local data-sharing scheme used Twitter to vent their opinions, instead of ignoring this or fighting

LESSONS LEARNED

What do these examples show us about the change in attitudes and culture needed to encourage greater information sharing that supports the needs and wants of service users?

- **Leadership:** Leadership is one of the biggest enablers of information sharing, especially in bringing together different organisations to operate, share information and respond as one. All the stories show the crucial contribution of strategic and clinical leaders to driving projects forward. Although representing different organisations, what these leaders had in common was an appetite for proactively engaging with local people; the willingness to answer questions directly; and the focus and commitment to deliver.
- **Building relationships and trust:** The creation and maintenance of strong, reciprocal relationships based on common values and outcomes was critical. It helped the different partners discuss issues constructively, secure 'buy-in' from patients and carers, and develop the knowledge and trust to articulate and agree a shared vision.
- **Communication:** Ensuring clear, simple, well-managed channels for two-way communications is key to allowing patients and other end-users to give their views.

back, Southend chose to engage directly to address them. The result was that the campaigner posted a positive video about the work and later asked how to promote it to other people with the same condition.

Tackling negative comments head on helped Southend turn a campaigning opponent of the scheme into an independent advocate – one of the most powerful aids to communication in public engagement.

SHOWCASE THE BENEFITS

The development of an IDCR in Bradford and Airedale is supporting health and social care integration and delivering improvements to a range of services. These include its end-of-life care, which reduces the number of unplanned hospital admissions that service users neither need nor want. Called 'Gold Line', the service provides a 24-hour telephone line for end-of-life patients, their carers and families. It allows patients' information to be shared and accessed so the best care can be given first time, in line with their wishes.

Angela, who is her husband's carer, sums up the positive impact of Gold Line:

"Since being part of Gold Line our lives are much calmer, and

it's been a real help knowing that there's always someone available to reassure or support me. I just wish we had had this service earlier, because it has taken the frustration and the fear out of what we have to do. One number, one call and everything is sorted - no chasing for appointments. I cannot praise the staff enough."

IF YOU ONLY REMEMBER ONE THING

Experience at national and local level tells us that information sharing can stand or fall on the effectiveness of public engagement. Information sharing and the development of the nuts and bolts for how it will work needs to be supported by effective and strong engagement and communication with the public. The prize is public trust, support and even championing of information sharing, because people understand, see and feel the benefits it delivers.

For more details about the IDCRs and the examples covered in this article visit www.informationsharing.org.uk/hsc.

Imogen Fuller

Engagement Manager,
Centre of Excellence for
Information Sharing

OPEN DATA FOR BETTER OUTCOMES

Data releases are powering new activity at Defra and attracting all kinds of users to explore and exploit their potential, as John Seglias, Chief Technology Officer, explains.

Within the next few years, a key ambition is for Defra's word association to be 'data', not 'badgers' or 'floods', reflecting a growing reputation for data-driven government and taking open data seriously.

So what has taken the department down that route?

Data has always been central to the services Defra provides, whether that is issuing flood warnings or monitoring wildlife. The 34 organisations that make up the Defra group hold a huge amount of data, from the location of bat-roosting sites, to 70 years of the Family Food Survey. Today, technological advances, coupled with high-powered computer processing capability, create an opportunity for the department's data to be used in ever more interesting and creative ways. In turn, this has generated the potential to really transform the way we undertake field activities.

However, this is not something government can do alone. Making it available to other people for a price could be an option, but the experience of many countries is that the paid-for data market is narrow. It also acts as a drag on the potential for data to act as an engine for economic growth. So, paradoxical as it may seem, the greatest value comes from making data freely available. This means not simply making the data available free of charge, but releasing it under a permissive licence that makes it easier to access, use and share across government, as well as in the private and not-for-profit sectors.

ASPIRATIONAL TARGET

With that in mind, in June 2015, Elizabeth Truss, then Secretary of State for Environment, Food & Rural Affairs, announced that Defra would release 8,000 data sets by

the end of June 2016. That target was in itself a departure from the norm for Defra. The 8,000 figure was not a bottom-up calculation based on detailed analysis of data holdings. It was an aspirational target designed to galvanise activity across the Defra group and allow others to unlock the potential of the data it held.

And it worked. The 10,000 data sets released by June this year – well above the target – have been put to great use by all sorts of users. The landmark release within these data sets was 11 terabytes of Lidar data (see box) in September 2015, and the list of sectors tapping into this resource is truly remarkable. British wine producers are using the data to help them decide where best to plant vines, identifying 'frost hollows' or badly drained areas that can affect their crop; architects are using it to build a model of London as they plan the next skyscraper; game developers to build new landscapes for Minecraft; and archaeologists to discover lost networks of Roman roads from Lancashire to Dorset. In October 2016 alone there were almost 21,000 downloads of Lidar data from data.gov.uk.

GAUGEMAP

Another great example of open data in practice is the GaugeMap API (a way for one computer to use information or services held on another computer, often across the internet). The GaugeMap API, developed by Shoothill, uses live river-level data from the Environment Agency and partner organisations in Wales, Scotland and Ireland. Information is updated every 15 minutes on Shoothill's website, a service that really came into its own during last winter's flooding. In a first for the 'Internet of Things' – the growing network of internet-connected objects able to collect and exchange data using embedded sensors – each of over 3,000 river-level gauges tweets its level twice a day. The companion FloodAlerts API provides the basis for apps from the Red Cross and even Facebook to keep people informed about flood risk. Despite these open data success stories, Defra recognises that the majority of those accessing its data are professionals. To be truly open, there is still work to be done to build tools that can open up the

Lidar is a laser-light equivalent of Radar. The Environment Agency (EA) has been using it since 1998, bouncing beams of light from planes flying over river catchments, coastal regions, water courses and surrounding land to produce 3D computer maps that help anticipate and plan for flood events, often using very powerful computers. The release of EA Lidar maps as open data in 2015, for free and with extremely permissive licensing, has meant that organisations in the private, public and not-for-profit sectors have been able to capitalise on the improved processor power and storage capacity of consumer-grade computers to access and process the data.

As well as more 'blocky' 1m and 2m resolution Lidar, perfect for importing into games like Minecraft, raw point cloud data made open in 2016 allows city planners and architects to build more realistic 3D environments. Variations in the reflected light points yields information about material attributes – whether it is the ground, a plant or water surface, for instance – which can be used to understand landscapes better. Now, academics and innovators are using data to gain actionable insights for use in insurance and facilities industries.



UK satellite programme

data sets to those of us without specialist IT skills.

Looking to the future, imagine living in a world where data captured by the network of satellites orbiting our planet are used to inform policy development and deliver cutting-edge services. Earth observation – the use of remote sensing data from satellites to gather information about the Earth’s physical, chemical and biological systems – is still in its infancy, but promises to be a valuable tool to increase the effectiveness and efficiency of decisions in a multitude of environmental fields including: crop health, forestry, flood management, habitat condition, water quality and coastal management.

VISION

Defra’s vision for Earth observation is “to ensure satellite data are playing to their full potential in policy development and operations across Defra by 2020”. The Earth Observation Centre of Excellence (EOCoE), a collection of Defra group organisations and external partners, including the devolved administrations and the UK Space Agency, has been set up to achieve this. The EOCoE shares data, information and expertise across the Defra group, and the programme has already proved valuable by identifying cost savings in delivery of the

Common Agricultural Policy and contributing to updates of the National Forest Inventory.

The potential for the EOCoE model of collaborative working to be rolled out across Whitehall is being explored under the UK Government Earth Observation Service and, continuing the open data theme, the Government’s environmental Earth observation data has now hit the world stage, having been made available through the international Earth observation ‘GEOSS’ portal.

While users and customers are the main beneficiaries of Defra’s focus on data, it has been positive for the department and its delivery bodies, too. The target to release 8,000 data sets generated energy and buzz across the Defra group. People didn’t spend lots of time debating the definition of a data set, instead they got on and released what was there, learning lessons as they went.

And while data has no monopoly on openness, people involved in data often seem to be leading the way in opening up the working culture. Defra’s extended data community (#defradatafunction) are role models for a more flexible, less hierarchical mode of working. They are hungry for improvement and thrive on trust. Inspired by their enthusiasm the department has embraced ‘un-conferences’, where topics are ‘pitched’ by participants on the day rather than being

planned meticulously in advance. Meanwhile the #defraselfie has gone viral, encouraging people to share more of themselves and contributing to a sense of connection across the Defra group.

A year or so ago the department developed a set of design principles for the Defra group:

- act together;
- make it simpler;
- avoid duplication;
- maximise impact; and
- embrace digital and data.

While only the last of these is specifically about data, it has become clear that the approach of putting data at the heart of everything Defra does meets all five of the design principles. Collaborating within and outside the Defra group, communicating straightforwardly, breaking out of silos and creating space for others to innovate is creating the organisation the department wants to be.

Ultimately this is about enabling the widest possible group of people to help deliver what matters most – positive outcomes. Open data, open ways of working and open culture will help make that happen.

You can read more about Defra’s approach to data on defradigital.blog.gov.uk.

John Seglias

Chief Technology Officer, Defra

BUILDING AN OPEN SOLAR POWER MAP

Michel Blancard, a data scientist at the French Government's Etalab, describes a data-led approach to expanding the use of solar power in France by creating a national register of suitable properties.

To open Civil Service Quarterly up to the widest range of innovative practices and experiences, we plan to showcase interesting case studies from governments around the world – of which this is the first.

Etalab, essentially the French Government equivalent of Government Digital Service, is the French Prime Minister's taskforce for open data, open government and data-driven transformation.

Despite the introduction of financial incentives for developing production of photovoltaic (solar) power systems since 2000, France ranks only 15th out of 28 in Europe for photovoltaic production per inhabitant.

As a comparison, Germany sets an example with production per inhabitant five times higher. Concerns about the selectiveness of the subsidies, and the increasing burden on finances, led to a reduction in the incentives after 2010.

While Germany and other countries developed solar cadastres (public registers of property) to assess the potential of candidate roofs for solar panel installations, such initiatives are still limited to a few cities in France, Brest and Paris being the most successful examples. These cadastres often use a three-dimensional model of a city, requiring expensive data

collection and treatments, and, consequently, are used mostly for highly populated areas.

An open solar cadastre, assessing the potential of roofs for solar panels covering the whole territory, would not only benefit public authorities but also a whole community comprising energy providers, panel installers, consulting companies and homeowners.

The Etalab team used an innovative, cost-efficient approach combining open data and open algorithms, relying on external contributions to build a nationwide solar cadastre.

The French land cadastre provides the contours of every structure. The shape of the roof is still uncertain and visual analysis is required to distinguish a ridge going west to east (suitable for solar installations) from one going north to south. So, satellite and aerial images



SEQ Illustration*ARABIC1:
Crowdsourcing platform, displaying a
roof image and 4 possible choices



covering the whole French territory with sufficient precision for most situations are used.

Etalab took advantage of a hackathon to design and set up a crowdsourcing platform with the help of enthusiastic developers. The platform displays the image of a roof and the user is invited to provide its orientation. The platform, being fun and somewhat addictive, received 100,000 contributions in a three-week span by word of mouth. This allowed us to classify 10,000 roofs with confidence. We identified just one case of vandalism, which was easily spotted and discarded.

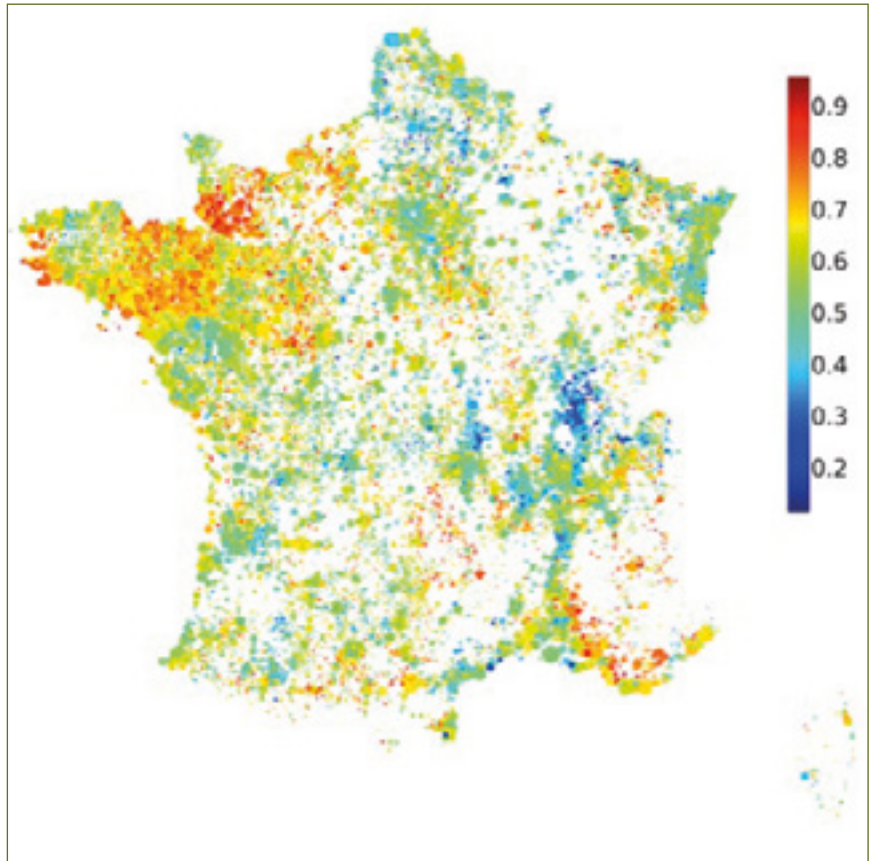
This is a small sample compared to the 50 million buildings in France, but it is enough to programme an automated classifier. Using standard techniques in image processing, namely logistic regression and deep neural networks, we obtained a classifier that was correct 80% of the time. Later, we found the automated results to be comparable to human contributions in accuracy. Run on standard hardware, the classifier takes one second to make a decision on an image.

This classification challenge was later taken on by a hundred teams during the Data Science Game, an international data science competition. The winning team, using newer neural networks and advanced techniques like data augmentation, fine tuning and ensemble learning, achieved a 30% lower error rate.

SOLUTION AND ACTION

The nationwide map of roof orientation shows differences between regions affected unequally by wind and topography. For example, most roofs in Brittany have a favourable orientation, whereas the opposite is the case in the Rhône Valley.

“ This map is of great importance to assess the relevance of solar incentives at a local level. ”



Roof orientation map. Regions with a high number of favourably orientated roofs are displayed in red. Source: SEQ Illustration*ARABIC2



Yearly sum of irradiation. Source: PVGIS © European Union, 2001-2012

It is worth comparing with the solar exposure to evaluate solar potential.

The solar cadastre doesn't take into account shades or the angle of inclination of roofs, and it discards roofs with complex shapes. However, it is intended to be completed by more precise, possibly local and expensive, data to deliver a better result. All the data², as well as the code³ used and produced in this project, is open and documented. Therefore, it is easy for anyone working in this field, whether from the public or private sector, to quickly build an improved solar cadastre on top of ours or to replicate it in another country.

LEARNINGS

As a government service attached to the Prime Minister, Etalab prefers inclusive approaches to closed ones. Machine learning projects like this one seem particularly suited to public participation. Every citizen, regardless of their skills and expertise, can offer their help using a crowdsourcing platform.

Calling on public participation is also a natural way to communicate and advertise the goals of the project.

The Etalab team benefited from voluntary contributions during the development of the crowdsourcing platform, the construction of the training set and the building of automated classifiers. Additionally, we rooted our project in open data and open source tools. In this way, a team of two people managed to develop this project within a few months, for negligible hardware costs. Although projects using machine learning are often termed 'Big Data', it would be a misnomer in this case, since we systematically favoured small-scale, quick and cost-effective methods, involving manageable amounts of data.

“ This approach lays the foundation for tools facilitating the work of public decision-makers involved in energy policies. ”

It could be easily replicated for similar issues; for example, detection of bus lanes and pedestrian crossings, for land-use

classification, and so on.

We demonstrated the application of state-of-the-art, free, open-source, well-packaged machine learning solutions outside of a research context. Engineers and developers can now extract value from images without having to be specialists in image processing or deep learning. It is probable that such tools will become increasingly widespread and eventually find their way into the general IT engineer's toolbox.

Michel Blancard
Data Scientist, Etalab,
French Government

FURTHER READING AND LINKS

- Šúri M., Huld T.A., Dunlop E.D., Ossenbrink H.A., 2007. Potential of solar electricity generation in the European Union member states and candidate countries. *Solar Energy*, 81, 1295-1305. <http://re.jrc.ec.europa.eu/pvgis/>.
 - Huld T., Müller R., Gambardella A., 2012. A new solar radiation database for estimating PV performance in Europe and Africa. *Solar Energy*, 86, 1803-1815.
2. <https://www.data.gouv.fr/fr/organizations/opensolarmap/>
 3. <https://github.com/opensolarmap>



DATA ANALYTICS FOR MORE EFFICIENT SERVICES AND BETTER LIVES

Andrew Goodman, Programme Director, Home Office Digital, Data & Technology Capabilities & Resources, looks at how the department is using data analytics to make people's lives better.



A great technology acceleration is under way in how, as citizens, we interact with and consume public services. At the same time the speed, efficiency, personalisation and security we demand of these services is increasing. Looking to 2020 and beyond, the forecast for a department like the Home Office is more: more people, more goods, more opportunity and more challenge.

International demographic shifts and population movements only serve to enhance the United Kingdom's long-established 'pull' factor on the global stage. Given that the Home Office's portfolio includes public protection, homeland security, borders, immigration and citizenship, identifying and delivering effective solutions that address these challenges head on is a clear priority. It is against this backdrop that data analytics can come into its own.

WHAT IS DATA ANALYTICS?

So, what is data analytics? Definitions vary, but it can be understood as the science of examining raw data with the intention of drawing conclusions about and from that information. As a discipline, it involves applying algorithmic or mechanical processes to uncover insights such as patterns or statistical correlations.

Many of us have passing familiarity with the role that data analytics and big data play in our lives, whether through commercial loyalty cards or targeted advertising based on our browsing or purchasing behaviour. Less, though, has been written about the transformative power of this technology for the delivery of high-quality public services, and it is time this changed.

For the Home Office, data analytics means the ability to

search across organisational data sets to assess the impact of existing policy and processes and – most intriguingly – to find previously unknown trends and connections. It might be used, for example, to predict trends, create statistical models on customer transactions, assess policy effectiveness, or provide specific data for operational teams to take forward investigations and enforcement action.

Strategically, perhaps the greatest value of data analytics is in its capacity to transform the effectiveness of an organisation by supporting credible, evidence-based decision-making on a scale and at a speed that was previously unachievable. It can provide a powerful lens for looking at your business, allowing you to question, probe, consider and shape how you deliver core services. Analytics isn't about looking for data that supports previously held conclusions. Neither is it about handing control or the decision-making process over to an algorithm, standing to one side and waiting for the computer to say 'yes' or 'no'. It's about drawing on and qualifying real human insight to identify more innovative and efficient approaches to how we work. It's about understanding where we have come from to better inform where we are going; listening to what our customers are telling us and delivering better, targeted services.

PROTECTING THE VULNERABLE

Analytics isn't a speculative or 'armchair' activity, it can play a direct and powerful role in the Home Office's work to protect the most vulnerable in society. Right

now, analytics is being used to support work to identify and catch the perpetrators and facilitators of modern slavery. Here, the role of analytics involves providing evidence to investigators and policy colleagues to understand the behaviours of perpetrators and victims; both in order to model and test the effectiveness of policy responses and to provide tactical insight to frontline colleagues. Here, analytics plays an increasingly important role as an enabler, supporting colleagues in identifying and adapting to changing circumstances on the ground and tailoring their interventions according to new insight and evidence.

‘ASSURANCE SCORING’

At the other end of the spectrum, analytics is being used to make life better for UK citizens through the use of ‘assurance scoring’. As the name suggests, this analytical service is primarily focused on identifying low risk or assured ‘good’ transactions. It works by looking at data attributes to identify clusters of variables that are typical of ‘good’ service transactions that the department’s operational experts have dealt with repeatedly in the past. Rule sets are then created that allow transactions to be sorted into categories representing different kinds of risk. Depending on what category of risk a transaction falls into it can then be directed more quickly to the relevant business area and resolved by colleagues with the appropriate expertise.

One area of Home Office business where the possible value of this capability is being explored is the passport application and renewal process. This is a high-volume customer service function that routinely deals with a wide range of products, customers and requirements that present different types of opportunity and risk. Most passport applications are legitimate, but some are from people who are not entitled to a UK passport. In some channels, effective use of assurance scoring could reduce the time, resource and layers of processing expended on interactions with those customers the organisational business already knows from its daily engagement

are likely to be genuine.

From a privacy perspective, use of assurance scoring by the Home Office may have the additional benefit of reducing the net ‘intrusion’ into the lives of many customers and citizens by reducing the level of data processing required to provide services to them. This might mean removing the requirement for face-to-face interviews, for example, or for additional supporting documentation in certain business work streams.

THE DATA ANALYTICS COMMUNITY

Used properly, data analytics offers a vital tool for testing, empirically and impartially, our own and the public’s preconceptions. This commitment to impartiality – a core tenet of the Civil Service – is being underwritten by the Analytical Quality Assessment (AQUA)⁴ process. This provides data analytics with a framework for quality-assuring statistical inputs, methodology and outputs against the risks their use represents.

The Home Office is not alone in its efforts to make the most of the potential of data analytics. Across government, centres of excellence are emerging, looking to better apply this capability to real-world challenges. There is a passionate and innovative data science culture that is looking to redefine how government delivers public services. Drawing this community together and providing a focal point is the Government Data Science Partnership (GDSP). The GDSP was created to help enable change and support this emergent network. It includes Government Digital Service, the Office for National Statistics and the Government Office for Science, and acts to promote the use of data science across government.

DATA SCIENCE ETHICS

One of the most valuable resources the GDSP has provided is a new Data Science Ethical Framework⁵. The framework seeks to help policy-makers and data scientists maximise the potential for data

science in government, while tackling the legal and ethical issues that accompany new forms of data analysis. Anyone with even a passing interest in the subject should read this document for a clear summary of the challenges that data scientists like those in the Home Office are grappling with on a daily basis.

Part and parcel of working in the world of data analytics and big data is responding to the many misconceptions that spring up around the potential uses of this capability. Perception is everything, and data analytics benefits enormously from being demystified. The Home Office, like the rest of government, takes its responsibilities as trusted data controller extremely seriously. This means not only ensuring that data is held securely and used in a measured and proportionate way to deliver functions and services, but ensuring that the public understands and accepts that this is the case.

As a May 2016 survey by Ipsos MORI⁶ for the GDSP showed, public approval for government using data analytics is actually quite high when used in this way. We have a responsibility to get this right. The challenge is to deliver services that are genuinely transformational but to maintain public approval for how we go about this. For the Home Office, this is at one with its core mission to keep citizens safe and the country secure. As threats change, the department needs to keep pace, responding to the opportunities of modern technology and providing value for money to taxpayers.

Andrew Goodman

Programme Director, Home Office Digital, Data & Technology Capabilities & Resources

4. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/416478/aqua_book_final_web.pdf
5. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/524298/Data_science_ethics_framework_v1.0_for_publication__1_.pdf
6. <https://www.ipsos-mori.com/Assets/Docs/Publications/data-science-ethics-in-government.pdf>

BUILDING AN INTERNATIONAL COALITION TO COMBAT ANTIMICROBIAL RESISTANCE

Professor Dame Sally Davies, Chief Medical Officer for England, Nicole Redhead, Global Health Policy Manager – International AMR, and Tracy Parker, UK AMR Strategy Programme Coordinator, describe how the UK has worked at home and abroad to galvanise coordinated action against AMR.

Antimicrobial resistance (AMR) is one of the biggest threats to health across the globe to emerge for many years.

The human and economic costs of AMR are already massive: the recent independent international review undertaken by Lord O'Neill estimated that if we do nothing to stop its spread, today's already large 700,000 deaths every year could become 10 million – that is more people than currently die from cancer. By 2050, the cost to the global economy will be around US\$100 trillion. We risk returning to the pre-antibiotic era, when infectious diseases were responsible for around 40% of all deaths, and where our inability to treat common infections with antibiotics will render many modern medical advances dangerous, including major surgery and cancer treatment.

However, AMR is not new. It has been understood for many years. Indeed, Alexander Fleming drew attention to the possibility of AMR when accepting his Nobel Prize for Medicine in 1945. There have been many attempts to address it. The first World Health Assembly resolution on AMR was agreed in 1998; the EU and the UK published strategies in the early 2000s. Yet, despite these efforts, resistance and, consequently, deaths were increasing, a problem exacerbated by the fact that there had been no new class of antibiotic brought into clinical use since the 1980s.

NO SIMPLE SOLUTION

It was clear that there was no simple solution and that the problem needed to be tackled on a number of fronts. The challenge is to stimulate action across a wide range of sectors, adopting a fully integrated multidisciplinary and cross-sectoral approach, for humans, animals, agriculture and the environment.

Using the role of Chief Medical Officer (CMO) for England has enabled us to become champions for AMR in the UK and across the world. The CMO annual report in 2013 called for action across all sectors, human and animal health, industry, academia and research funders. That call was reinforced in a comprehensive cross-government UK AMR Strategy published in September 2013.

The UK continues to champion AMR and the CMO is a figurehead. In practice, implementation of AMR strategy in the UK has been overseen and driven by an interdepartmental steering group. The Department of Health (DH) chairs that group and coordinates the approach with key partners, Public Health England (PHE), the Department for the Environment, Food & Rural Affairs (Defra), NHS England and the Devolved Administrations. The lead partners, in turn, work with other departments and agencies, professional bodies, industry and the research community.

Coordinating delivery of the UK strategy across such a wide range of organisations has not

been without its challenges. Early on, when the detailed implementation plan was in development, many health sector agencies were newly established and still finding their feet. Partner departments and agencies have different approaches to project management, different priorities, stakeholder groups and levels of resourcing. There was no new funding for implementation when the strategy was published and departments have had to determine how to deliver their commitments within existing funds. The DH team has had to navigate sometimes choppy seas where DH is ultimately responsible for delivery of the programme on behalf of the Government, while needing to acknowledge the prerogatives of its partner departments and agencies.

An additional complication is the fact that, although this is a UK strategy, each of the Devolved Administrations has responsibility for implementation within their own jurisdictions in relation to human health, while Defra has UK-wide responsibilities in relation to animal health and the environment.

In addition to these challenges, the evidence in relation to resistance, its growth and transmission has continued to develop and the implementation plan has needed to continuously evolve in response.

Progress includes a reduction in the prescribing of antibiotics for humans in England by 4.3% between 2014 and 2015, and

a reduction in the sales of antibiotics in the animal health sector in the UK by 9% from 2014 to 2015. This country now has one of the most comprehensive surveillance systems in the world to identify problems early and take action. There are unprecedented levels of research collaboration and increased investment.

In the last year, PHE has overseen the development of a data-sharing website called 'Fingertips'. This makes relevant data easily accessible to clinicians and the public and allows local areas to compare themselves to other similar areas. The Fingertips data portal is crucial in supporting local teams to deliver the new ambitions for the reduction in inappropriate prescribing and gram-negative infections and the use of antibiotics in animals, announced earlier this year. Gram-negative bacteria are a particular concern because they are resistant to multiple drugs and are increasingly resistant to most available antibiotics. They can cause pneumoniae, bloodstream infections, surgical site infections and meningitis.

Of course, one of the biggest challenges domestically is to change behaviours: a third of the public believe that antibiotics will treat coughs and colds; and one in five expect to be prescribed antibiotics when they visit their doctor. Work will continue at national level to raise awareness, and at local level to support doctors to improve their prescribing.

LEADING THE WAY

All of this domestic work would be of marginal significance if AMR was not addressed globally, and that has been the greatest challenge. There is recognition that while all countries need to act, many will have other priorities and low levels of capacity and capability. However, the UK has led the way, building international coalitions, raising awareness, and encouraging and supporting other countries to take action.

One key message was important from the start: the need for access not excess. This recognises that, in many countries, lack of access to antimicrobials still kills more people

than resistance. Accordingly, UK funding initiatives have focused on low- and middle-income countries to help develop their capability and capacity to respond.

“ There is no doubt that much of the success to date, both domestically and internationally, is the result of lead departments working with experts, each other and across sectors to develop and implement policy. ”

To galvanise global action, collaboration with other government departments was broadened, setting up a cross-Whitehall group chaired by the Foreign & Commonwealth Office (FCO), and bringing together partners from the FCO, DFID, Treasury and Cabinet Office, as well as DH and Defra. That group had clear and ambitious objectives – so ambitious many doubted it was even possible. Every opportunity was taken to promote AMR with governments in other countries and international bodies through international networks. This included work with like-minded countries to amplify key messages and using global partnerships like the Global Health Security Agenda to build consensus with other nations. The aim was to strike a balance between being perceived as telling other countries what to do and empowering them to take action.

In May 2014, with partner countries, the UK secured a World Health Organisation resolution requiring member states to develop national plans to address antibiotic resistance. This was followed a year later by a Global Action Plan (GAP). That plan gives the world the blueprint for action needed to tackle AMR. However, implementation of the GAP requires much greater cross-sector and international collaboration. Key to achieving this is gaining

political support at the highest level internationally. Making the case that AMR has social and economic consequences that have to be addressed is crucial to delivering the UN's Sustainable Development Goals.

The period since GAP has seen a series of international communiques, declarations and strategies, culminating in the G20 and UN General Assembly (UNGA) in September 2016. At G20, the UK secured commitments to take forward work to explore how to mitigate the spread of resistance and promote further research and development into new antimicrobials. At UNGA, 193 countries agreed an ambitious political declaration to tackle AMR. This was only the fourth time that a health topic has been discussed at the General Assembly and, as such, is a significant achievement.

The challenge now is to turn the commitments made under the UN declaration and the G20 agreement into international action. There will now be a pause to allow the UN organisations to look at the options, while continuing to influence through and with the G7, G20, WHO and the UN General Assembly.

There is no doubt that much of the success to date, both domestically and internationally, is the result of lead departments working with experts, each other and across sectors to develop and implement policy. The UK has demonstrated its ability to move quickly to establish cross-government teams focused on specific initiatives and objectives, and to support senior officials, ministers and the CMO to galvanise the support needed to make a difference to this international problem. The challenge now is to sustain this momentum, recognising that we, and the world, still have a long way to go to tackle the global threat of AMR.

Professor Dame Sally Davies - Chief Medical Officer for England,
Nicole Redhead - Global Health Policy Manager – International AMR, and **Tracy Parker** - UK AMR Strategy Programme Coordinator

7. [https://amr-review.org/sites/default/files/160525_Final paper_with cover.pdf](https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf)

DATA IN THE LIFE OF...

We interviewed two data specialists about the uses of data and data science, and what it's like to be a data scientist in government.



Charlie Boundy

Charlie Boundy is Head of Data Science in the DWP.

WHY IS DATA AND THE USE IT CAN BE PUT TO IN GOVERNMENT SO IMPORTANT?

Data may sound a bit dull to some, but it's what makes the world go round right now. Whether you're aware of it or not, the phone in your pocket, the computer on your desk, and the increasing number of connected devices we take for granted, are busy generating more data than ever before.

But it's the way data is being used that is rapidly evolving. Big companies have made data their product. They use sophisticated analytics to display only the most relevant results based on what they believe you or 'people like you' are interested in. They may not always get it right, but it all makes everyday decisions quicker and more convenient. And it's all based on what they've learnt from millions of their users.

So, in government, when we're tackling policy issues or delivering citizen-facing services, why can't we tap into the experience of potentially millions of users? In everyday life we ask ourselves questions like: How big a problem is this? Is it normal? And what's the best possible outcome? All of us look for reassurance that there's evidence to back up our decisions, but usually we have to rely on instinct and the limits of the information provided.

Public expectations have gone way beyond just being able to access services online. Data is central to the next wave of digital transformation of public services, and that means getting the right data to the right place at the right time.

SO, IF THIS IS THE NEXT WAVE OF 'DIGITAL TRANSFORMATION' THERE ARE BOUND TO BE LOTS OF BUZZWORDS. WHICH ONES DO WE NEED TO KNOW?

I'll give you my interpretation of three particular ones that are

increasingly becoming part of common language.

First, '**Big Data**'. This refers to the wider data captured about a digital transaction, event or service. For instance, knowing the type of device used when someone was online, the 'journey' they took through your services, and how long they took, are all essential raw material to understand how effective a service is.

Secondly, '**data science**' refers to a set of skills and tools to explore data in order to flush out any trending behaviours or predict the likelihood of various outcomes. Personally, I believe there are two parts to this. One is using various analytics techniques on the data to discover a trend or pattern; the other is testing how that insight can be applied to a new or improved end 'product'.

One set of techniques used by data scientists is called '**machine learning**'. This is a valuable way of making automated decisions or recommendations more intelligent by learning from the outcomes of hundreds of daily interactions. Predictive text on mobile phones is an everyday example of this.

HOW ARE YOU APPLYING DATA SCIENCE IN DWP?

The first step is deciding where to start. DWP is a large, complex public service organisation where there are thousands of unique daily decisions affecting millions of people. However, as with all things digital, it starts with the user. So, we're looking at specific user groups, such as policy-makers and work coaches, and asking, 'What do you need to know when you are doing your job?'

Using that 'information need', it's then a matter of finding the right data from within the digital footprint left behind by thousands of others. It's not always there to be found but it's vital to use real data in any analytics to see what it's telling us.

The next step is translating that data into something usable or 'actionable'. It's no use offering raw data to a user who has seconds to make a decision but would need hours to interpret it every time they need it. Good data science needs to combine with good design skills to hide all that complexity with a simple set of insights or options for a user.

A great example is our work with policy-makers. We've been working on a data visualisation product that allows policy-makers to explore key statistics and understand more about significant geographies, trends and demographics. This helps them answer for themselves the most frequently asked questions, such as understanding how the volumes of claimants are changing over time in different areas. In fact, this was based on open published data. But that data in its raw form isn't easy for a non-technical user to get answers from.

So, by embedding analytics into apps or business processing, we want to help users such as work coaches and policy-makers understand which claimant groups need the most support, and customise our efforts accordingly.

SO WHEN WILL THESE SORTS OF TOOLS BE WIDELY AVAILABLE AND WHAT ARE THE CHALLENGES?

In ten years' time, you'll be hard pushed to find any systems in UK businesses that don't include some form of embedded analytics.

Achieving this sort of embedded analytics will transform any department, but there are three main challenges to overcome. The first is about having trusted data.

Data collection in most public services was originally

intended to enable systems to confirm customers' entitlement or liability on things like benefit payments, passport applications or tax payments. That data was soon being used for measuring operational performance and designing policy. Beyond that, any new ways of using data need to ensure that the quality of the data is reliable.

The next challenge is one of design - that is, creating products for people who don't like data or feel uncomfortable using digital services. Good design is vital to communicating data. Presenting too much data is overwhelming, and too little leads back to mistrust. How much depends on the user and the time they have to make their decision. In any situation, it's about shortening the time it takes to understand the details and assess the options.

The final challenge I'd emphasise is about privacy and ethics. In other words, just because, technically, you can, doesn't mean you should.

Using big data, data science and machine learning opens up a wealth of possibilities, all of which sit within legal frameworks. But privacy needs to be designed in and, even then, sometimes, a solution may not feel right ethically. So, it's important to test ideas out first with users and think through the potential for unintended consequences. For example, would you welcome a text message for a discount offer to a coffee shop you are about to walk past? That seems a bit creepy to me. The Government Digital Service has published an ethical framework for data science, which is a good set of guidance to go by and build on.

WHAT DOES YOUR ROLE AS A HEAD OF DATA SCIENCE INVOLVE ON A DAY-TO-DAY BASIS?

There's definitely no typical day. In this first year it's been a mix of helping stakeholders understand the possibilities for data science and exploring a range of projects that test

different scenarios. Meanwhile, we've been building up a team with the necessary skills, as well as the technology platform to work on.

Some aspects have worked well, some haven't, but there's been a fantastic drive and collaboration which I've really enjoyed being part of. We've achieved a lot and learnt a lot.

The year ahead is all about realising the benefits of the projects where we can move more quickly, and maintaining our persistence in those we know will take longer.

HOW ARE YOU GOING ABOUT ATTRACTING MORE DATA SCIENTISTS AND DEVELOPING THEM IN THE DEPARTMENT?

Data science skills are in high demand across all industries, but few organisations have as meaningful a challenge as DWP. For some, the opportunity is a great chance to develop their skills, and for others it's a chance to work with a rich set of data that will grow exponentially with digital services. However, almost universally, the chance in government to make a difference to people's lives exceeds that of other organisations, which may be using data science in retail, manufacturing or something similar.

WHAT DOES THE FUTURE HOLD FOR DATA-LED SERVICES?

Making better use of data is more achievable than ever before given the amount of free software, cloud storage and collaborative communities out there. One of the best ways to start getting the most out of your data is in an offline 'sandbox' environment, bringing snapshots of 'safe' data together and trying out different data science and visualisation techniques.

Being truly transformative, though, means working closely with users to help them explore what data products they would find valuable and tackling specific questions, rather than trying to tackle them all.



Aimee Murphy is a data scientist at the Home Office. She was in the first cohort of the Data Science Accelerator programme for government analysts looking to develop their data science skills.

WHAT DOES DATA SCIENCE MEAN IN THE HOME OFFICE AND WHY IS IT IMPORTANT

Data science, or 'data analytics' in the Home Office means gaining new insights and analysis from our data and delivering services to Home Office operational areas that make their jobs faster, easier and safer. This could be through automating the finding of relevant information, building machine learning models to stream applications, or spotting potential victims of modern slavery.

For me, the modern slavery example is really important, because it shows the real and immediate benefit that use of data science can have in people's lives.

Obviously, I can't go into exact detail about how this is achieved, but it involves understanding the behaviours of the perpetrators and victims of this awful crime in order to pre-empt and disrupt future activity and uncover those who attempt to mask their crimes.

The team are immensely proud to be making a difference in this field.

Successful delivery, though, doesn't just require data

scientists. We work with other professionals in the Home Office Data Analytics Competency Centre including business analysts, infrastructure managers, user researchers and developers.

WHAT MAKES A GOOD DATA SCIENTIST

This is quite a point of contention. What makes a good data scientist depends on the project, the context and the environment they're working in. As a minimum, a data scientist should have an understanding of predictive analytics techniques, software engineering skills and be able to communicate in a clear, understandable way.

WHY DID YOU JOIN THE CIVIL SERVICE?

My first role in the Civil Service was as a temp when I was 17, preparing architectural documents for scanning. I joined the Civil Service permanently as a statistical officer in 2012. I was attracted by the idea of using data and analysis to do good – supporting evidence-based policy-making and informing the public through publishing open and accessible statistics.

WHAT ATTRACTED YOU TO DATA SCIENCE?

Being able to analyse and gain new insights from unstructured data, free text, images, metadata, and so on. The constant flow of new tools and techniques being developed

- the expanding art of the possible - is also very exciting.

HOW HAVE YOUR ROLE AND YOUR SKILLS DEVELOPED IN THE CIVIL SERVICE?

I took part in the first cohort of the Data Science Accelerator programme working on a project to web-scrape, analyse and interactively present road condition data.

The Accelerator programme is a cross-government training programme backed by GDS, ONS, GO Science and the analytical professions. Participants spend three months working part-time on a data science project from their home department with the support of GDS and a dedicated mentor.

The programme introduced me to some data science techniques and programming, but also taught me how to teach myself new techniques, how to find information and use (and contribute to) open source resources.

Since taking part in the Accelerator, I've worked in data science roles at the Department for Education and now the Home Office, and become a mentor on the programme myself.

WHAT HAVE BEEN THE POSITIVES FOR YOU PERSONALLY OF DEVELOPING THESE NEW SKILLS? AND WHAT HAVE BEEN THE CHALLENGES?

I've really enjoyed learning new techniques and programming languages. Broadening the range of techniques I can apply to a problem has made my work more varied and allowed me to generate new insights from data.

WHAT DO YOU THINK THE CIVIL SERVICE/GOVERNMENT NEEDS TO DO TO REALLY EMBRACE DATA SCIENCE?

Ensure we have the tools to do our jobs effectively – infrastructure, software, development time.

HOW DO YOU SEE YOUR FUTURE IN GOVERNMENT?

Data science is a developing area with new roles and avenues of research opening all the time. I'm excited to be a part of it.

TRANSPARENCY AND EVIDENCE – SHOW YOUR WORKINGS

A personal view on the use of data in government from Dr David Halpern, Chief Executive of the Behavioural Insights Team (BIT). BIT, a social purpose company, uses behavioural science insights to help make public services more cost-effective and easier to use, and to encourage people to make better choices for themselves and society.



Dr David Halpern

As well as using data and analysis, it's important to ensure that it is open to others to check and interrogate. In the academic world, many journals now publish the underlying data behind papers so that other researchers can check the analysis and consider if there might be other interpretations. It builds confidence in the results, and sometimes leads to important new perspectives.

Public bodies should do the same. This government has committed to be the “most transparent government in the world”, with transparency, openness and accountability embedded in everything it does. This includes not just data in the form of numbers, but also the evidence on which policy is based.

The 2013 review of the Civil Service Reform Plan committed to publish more supporting

evidence and data for policy. This was reinforced in the Open Government Action Plan 2016-18, setting out changes to publishing infrastructure to support this. Practical initiatives are putting clear evidence into the hands of public service professionals and commissioners, such as the ‘What Works’ evidence clearing-houses. The recent announcement of changes to the budget are similarly intended to provide more opportunity to explore the data and evidence behind tax and budget proposals.

Pressure is building to make sure that Whitehall ‘walks the talk’. In 2015, the Institute for Government (IfG) developed a rapid assessment tool to rate government departments on the use of evidence behind policy decisions. The Show Your Workings tool assesses how well government publications set out:

- a clear diagnosis of a policy problem;
- the evidence behind proposals to address it;
- evidence of value for money and an implementation plan; and
- how the policy will be evaluated – especially where the underlying evidence is limited – and then adapted.

In 2016, Sense about Science worked with the IfG and the Alliance for Useful Evidence, with cooperation from government directors of analysis, to apply the framework to a cross-section of policies. They published a report, *Transparency of Evidence*, in late November, highlighting many examples of good practice, and recommendations based on the

difference between the best and worst examples they found.

Next year’s report will go further. Sense about Science are now assessing policies from July 2016 – 2017 and planning to produce a departmental ranking next November.

Events are planned to support analysts and policymakers to understand how the scoring will work and how they might take action now to influence their ranking. Government publications can be let down by things that are easy to fix, such as failing to reference sources or provide the data on which claims are based. Sometimes the issues require more thought, such as designing the way policy is implemented so that it can be robustly evaluated and then adapted and improved as necessary.

This greater scrutiny provides extra impetus for departments to lay out the data and analysis behind policy announcements. But the real reason we should do this is because it’s good government. Laying out our evidence and data enables others to check our reasoning and to bring in further evidence and perspectives. It helps the public to engage and experts to assist and challenge us. In short, it strengthens the foundations of our policy and of democracy itself.

Dr David Halpern

Chief Executive of the Behavioural Insights Team (BIT)

LINKS

Behavioural Insights Team
<http://www.behaviouralinsights.co.uk>

OGI

© Crown copyright 2017

You may re-use this information (excluding images and logos) free of charge in any format or medium, under the terms of the Open Government Licence.

To view this licence, visit www.nationalarchives.gov.uk/doc/open-government-licence/version/2/ or email psi@nationalarchives.gsi.gov.uk