



**Data Science  
Campus**

# **Data Science is a team sport**

**Building capability through  
partnerships & knowledge exchange**

**Tom Smith**

**Director, Data Science Campus UK**

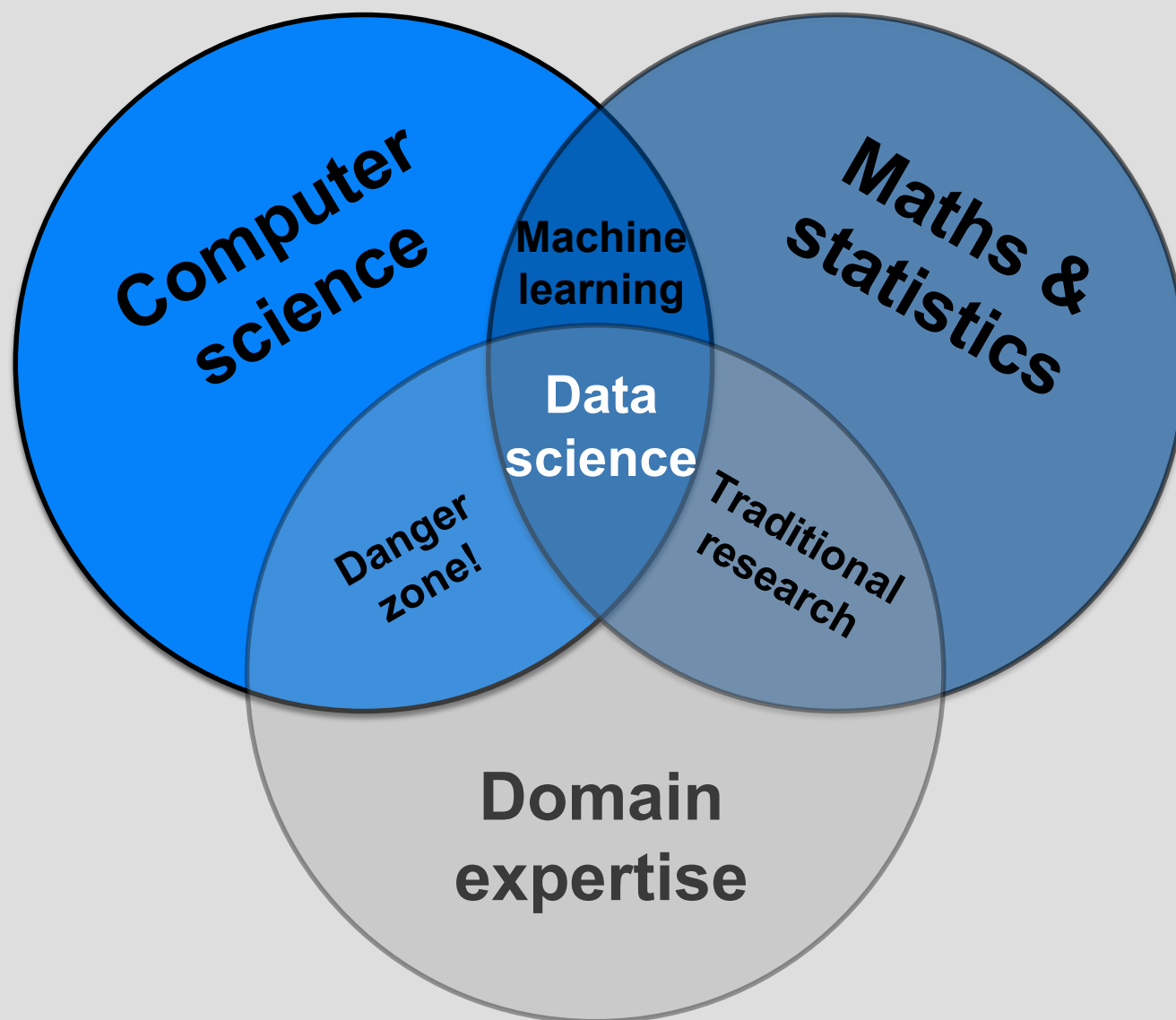
**data science**



“The 21st Century has brought new challenges in the analysis of data, and it is increasingly apparent that solutions to these are both statistical and computational. This has led to a great demand for people both in industry and in research who are able to draw upon the mathematics of both computation and probability to make sense of the large amounts of data that are collected in order to solve major problems.

Data science is an interdisciplinary response to this demand”

- University of Warwick



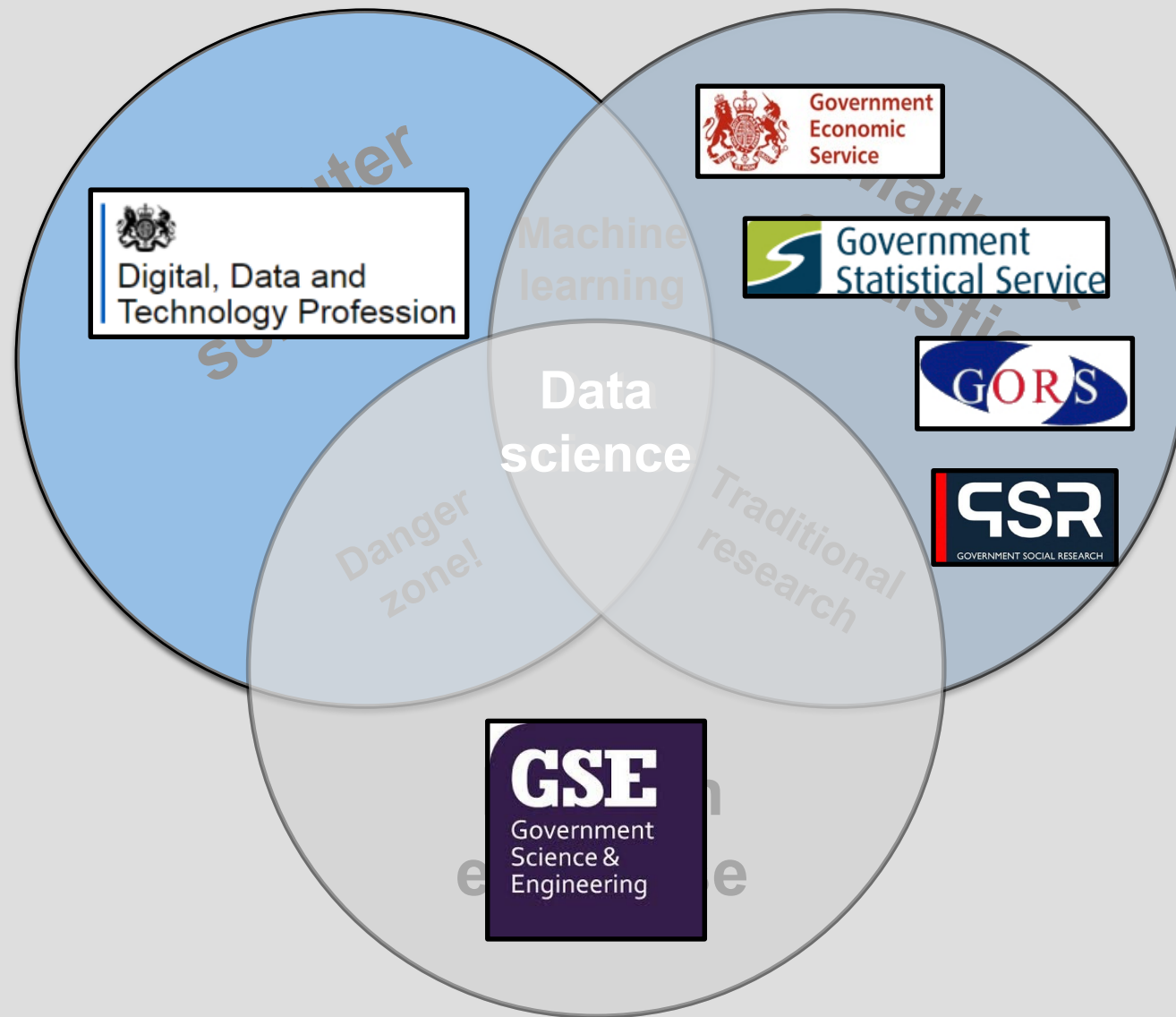
The Data Science Venn Diagram, designed by Drew Conway



“The 21st Century has brought new challenges in the analysis of data, and it is increasingly apparent that solutions to these are both statistical and computational. This has led to a great demand for people both in industry and in research who are able to draw upon the mathematics of both computation and probability to make sense of the large amounts of data that are collected in order to solve major problems.

Data science is an interdisciplinary response to this demand”

- University of Warwick



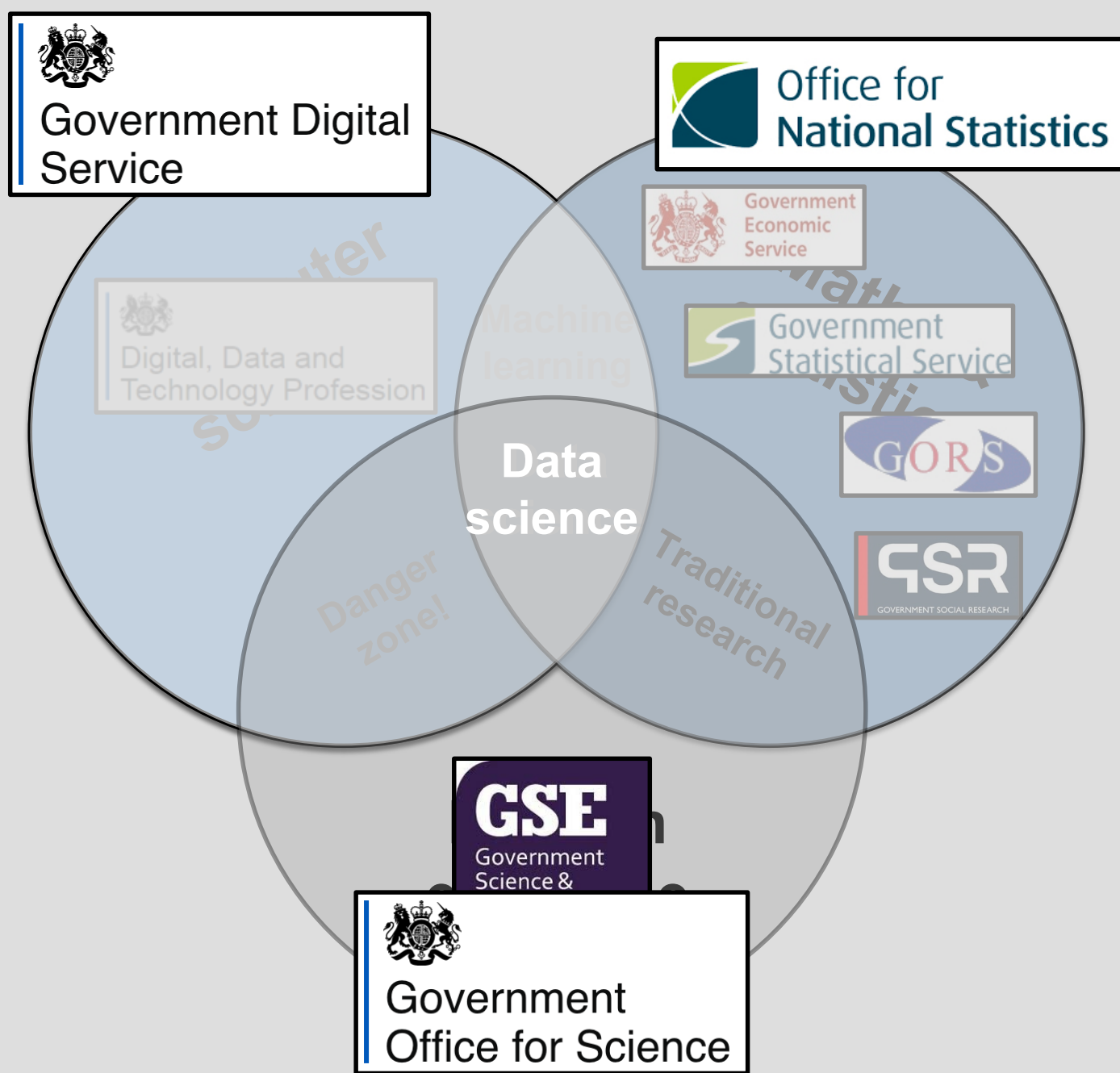
The Data Science Venn Diagram, designed by HMG



“The 21st Century has brought new challenges in the analysis of data, and it is increasingly apparent that solutions to these are both statistical and computational. This has led to a great demand for people both in industry and in research who are able to draw upon the mathematics of both computation and probability to make sense of the large amounts of data that are collected in order to solve major problems.

Data science is an interdisciplinary response to this demand”

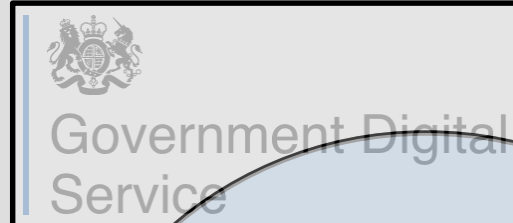
- University of Warwick





“Data scientists solve complex business problems using a combination of domain expertise, coding knowledge, machine learning and statistics skills on large and varied datasets.”

- Government Data Science Partnership

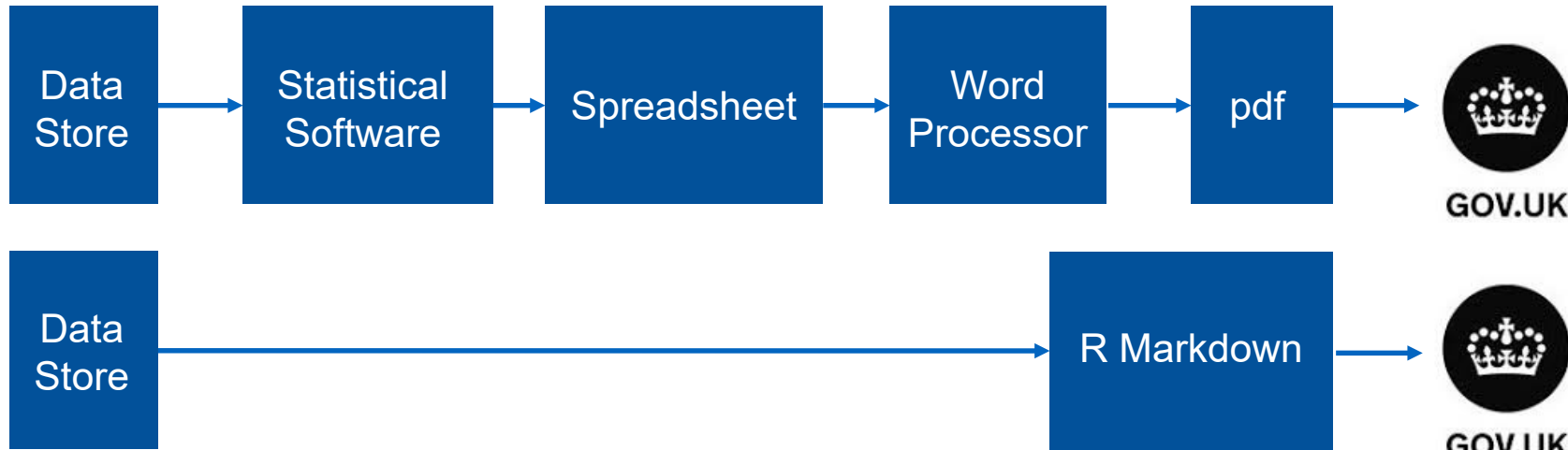


# Government Data Science Partnership

Cabinet Office | Government Office for Science | Office for National Statistics



# Reproducible Analytical Pipeline



## Efficiency Savings

**£118M**

Estimated annual efficiency savings across government stats publications

**£8.8K**

Estimated average annual saving per publication

### The Challenge

- Producing official statistics for publications is a key problem: as it is a time consuming meticulous process
- It is time consuming as the analysis has to pass through multiple systems and multiple individuals
- The systems are diverse and do not always conform to good software engineering practice

### Solution

- Use of software engineering tools and techniques such as version control.
- Automated generation of tables/charts and statistical verification
- Process from data storage to report generation

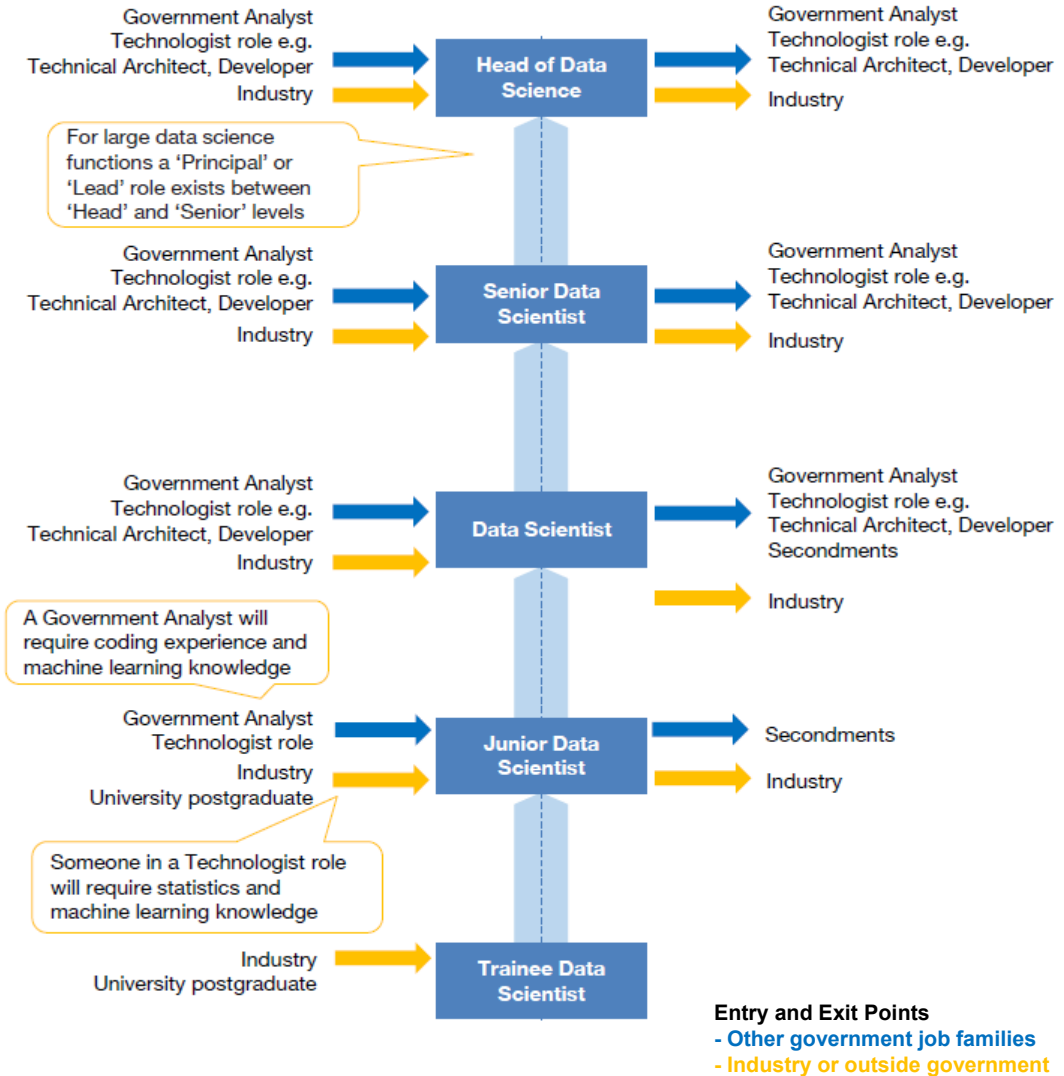
GDS estimate based on (i) c.16k statistical publications across government <https://www.gov.uk/government/statistics>; and (ii) the annual monetised efficiency saving or gains of learning RAP techniques in R or other programming language based on sample case studies

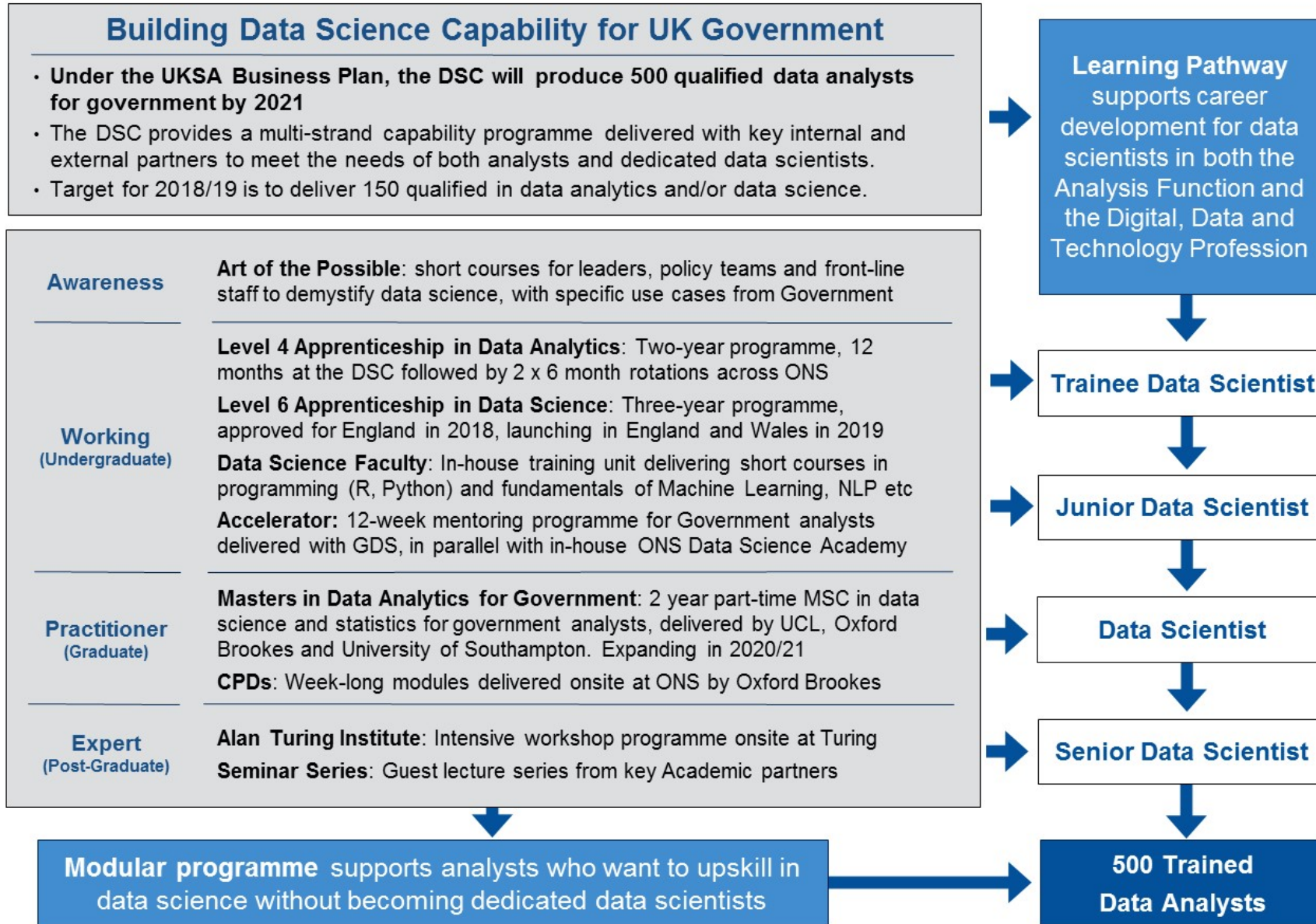
# Career Pathway

## A roadmap for a data science career in Government



Role Level	Description of role level and common activities
<b>Head of Data Science / Lead Data Scientist</b>	These roles provide leadership and direction across a programme of multidisciplinary data science projects, managing resources to ensure delivery. They are recognised as a strategic authority with technical expertise in cutting-edge techniques, defining the organisation's vision. They are a role model to other Data Scientists and champion adoption of best practice. They communicate with senior stakeholders and convince them of the strategic value of data science. They are a champion for the use of data science across government.
<b>Senior Data Scientist</b>	Senior Data Scientists are experienced Data Scientists who provide support and guidance to teams. They are a recognised authority on a number of data science specialisms within government, with some knowledge of cutting-edge techniques. They may work on projects of high political exposure, value or complexity. They engage with senior stakeholders and champion the value of data science. They line manage more junior colleagues. They communicate the value of data science to senior stakeholders.
<b>Data Scientist</b>	Data Scientists are proficient in data science. They have recognised technical ability in a number of data science specialisms and provide detailed technical advice on their area of expertise. They draw on other technical and analytical standards from across government and industry. They promote and present data science work both within and outside of the organisation. They engage with stakeholders to demonstrate the value of data science and propagate data science skills in other teams. They line manage and mentor Junior Data Scientists and manage small project teams.
<b>Junior / Associate Data Scientist</b>	Junior or Associate Data Scientists are responsible for aspects of existing data science projects, whilst gaining valuable hands-on experience. They are able to apply certain data science techniques and work to develop their technical ability. They adhere to the data science ethics framework. They work as part of a multidisciplinary team with Data Architects, Data Engineers, Analysts and others and provide limited advice on data science projects within teams. They identify and communicate lessons learnt during projects and follow good practice. They clearly communicate the value of data science work to stakeholders.
<b>Trainee Data Scientist/ Apprentice</b>	Trainee Data Scientists and Apprentices are given experience of practical data science work under supervision from more senior colleagues. They move from a strong awareness of core data science skills of coding, machine learning and statistics to a more effective working knowledge and develop their understanding of how to apply data science to business problems.







### **Level 4 Apprenticeship in Data Analytics:**

Two-year programme, 12 months at the Campus followed by 2 x 6 month rotations across ONS

### **Level 6 Apprenticeship in Data Science:**

Three-year programme, approved for England in 2018, launching in England and Wales in 2019

### **Data Science Faculty:**

In-house training unit delivering operational courses in programming (R, Python) and fundamentals of Machine Learning, NLP etc

### **Accelerator:**

12-week mentoring programme for Government analysts delivered with GDS, in parallel with in-house ONS Data Science Academy

### **Masters in Data Analytics for Government:**

2 year part-time MSc in Data Science and Statistics for government analysts, delivered by UCL, Oxford Brookes and University of Southampton. Expanding in 2020/21



Elund Morgan, Minister for Welsh Language and Lifelong Learning and Alexis Fernquest, ONS Data Analytics Apprentice, Nov 2018



## Attracting talent tomorrow, Accessing skills today

### University research programmes and PhD funding:

Research partnerships with over a dozen universities and the Alan Turing Institute, which additionally hosts the *Data Science for Public Good* joint PhD programme

### Centres for Doctoral Training:

7 UKRI, EPSRC and STFC centres for Doctoral Training in AI, Data Science, Statistics and Physics, with PhD co-supervision and Campus placements

### Graduate Student Placement Programme:

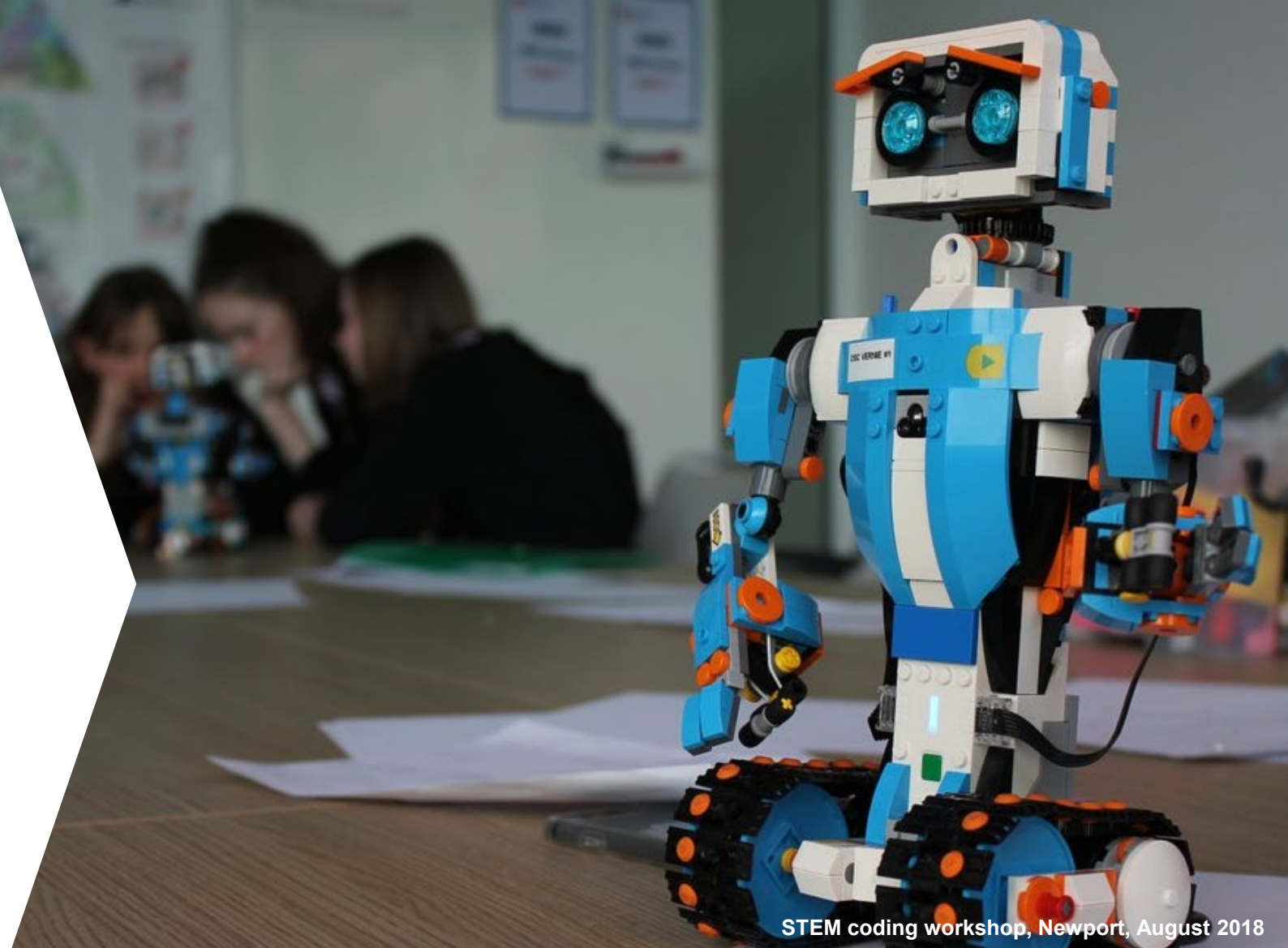
Three-month MSc thesis and MSc/PhD industry placements at the Campus

### Welsh Data Science Graduate Programme:

2 year MSc programme with Welsh Contact Centre Forum that includes three eight month industry placements.

### STEM Outreach:

Girls into STEM (Science, Technology, Engineering & Maths) programme for GCSE students; Nuffield Research Placements for A-level students

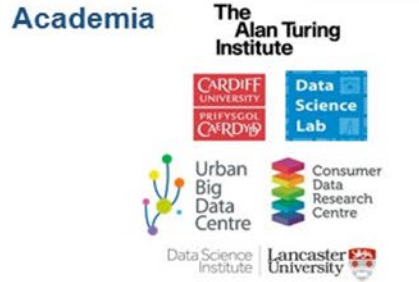


STEM coding workshop, Newport, August 2018

## Developing the next generation of Data Scientists

# Partnerships

## DSC partnerships facilitate access to data, skills and resources



**Joint Research Programmes** range from specific projects with individual universities (WBS Data Science Lab) to the Turing-HSBC-ONS Economic Data Science Awards

**Data Science for Public Good PhD Programme at the Alan Turing Institute** launches in Sep with four co-funded PhD students, and a 6 month placement at DSC

**Sunsetting PhD Co-funding Programmes** in place at Cardiff and Lancaster

**Graduate Student Placement Programme** at the DSC to undertake MSc thesis work (Southampton and Cardiff Met), with PhD placements launching in Sep (UCL)

**Centres for Doctoral Training** will become a strategic focus from 2019/20. The DSC is currently supporting 13 UKRI and EPSRC bids in both AI and Statistics



**Industry Collaboration Programmes** focused on public good outputs with key partners (PwC, Barclaycard) include data sharing, secondments, apprentice development, knowledge sharing events etc

**Skills and Standards** collaborations with industry working groups led by Royal Statistical Society, the Royal Society, The Alan Turing Institute, and the ONS-led Level 6 Apprenticeship Trailblazer group



**Government Data Science Partnership** with the Government Digital Service and the Government Office for Science drives growth of data science across government through training (Accelerator), knowledge sharing events (Government Data Science Conference, Community of Interest) and communications (Slack)

**Joint Training Programmes** with GSS Learning Academy and GDS Academy

**Direct Collaborations** with individual government departments to provide training and undertake joint research

**Data Enabled Change Accelerator Programme** delivered with the Department for Digital, Culture, Media and Sport under John Manzonei's Data Advisory Board provides DSC mentoring for strategic data science projects



**Research Partnerships** with international centres of data science excellence including the Centre for Big Data Statistics at Statistics Netherlands, and the United Nations Global Pulse Lab in Jakarta

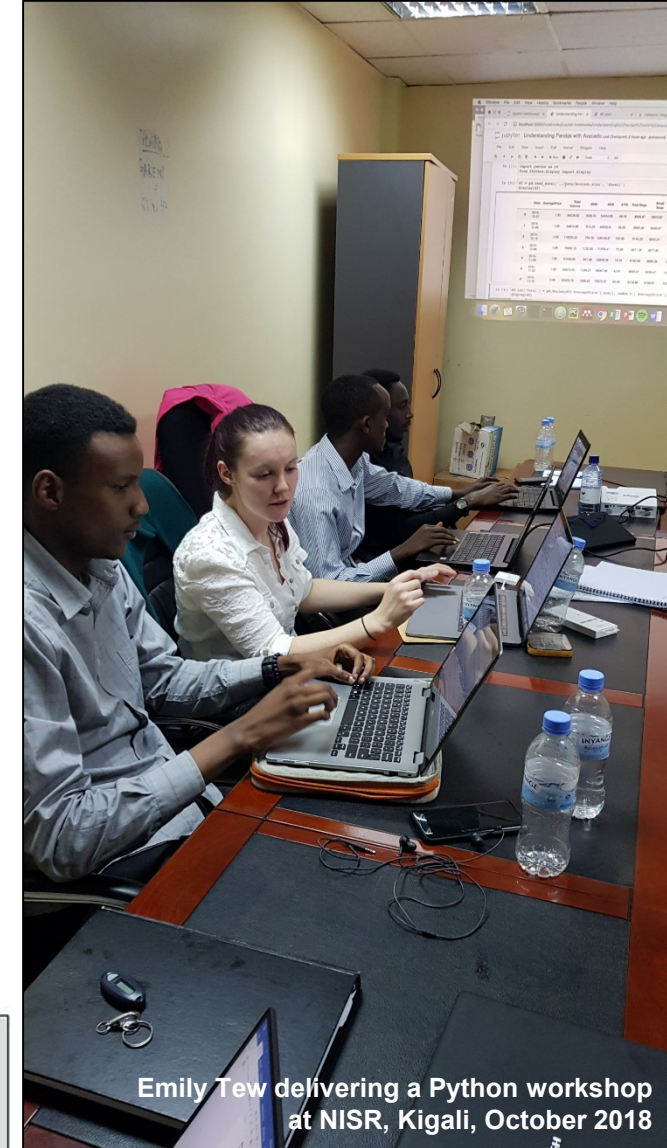
**Capacity Building Programmes** that support the development of public sector data science and data analytics skills in less developed countries, with a flagship programme with National Institute of Statistics Rwanda

**Learn:** Act as a conduit for new data sets, technologies and methodologies to enter ONS

**Grow:** Support the growth of data science across the UK and internationally

**Train:** Develop the next generation of data scientists and analysts in Government

**Recruit:** Promote ONS and Government as first destination of choice for graduates



Emily Tew delivering a Python workshop at NISR, Kigali, October 2018



Better Data Science for...  
...Better Decisions

# Thank You!

ONS Data Science Campus

Email: [datasciencecampus@ons.gov.uk](mailto:datasciencecampus@ons.gov.uk)

Web: [datasciencecampus.ons.gov.uk](http://datasciencecampus.ons.gov.uk)

GitHub: [datasciencecampus.github.io](https://github.com/datasciencecampus)

Twitter: [@DataSciCampus](https://twitter.com/DataSciCampus)

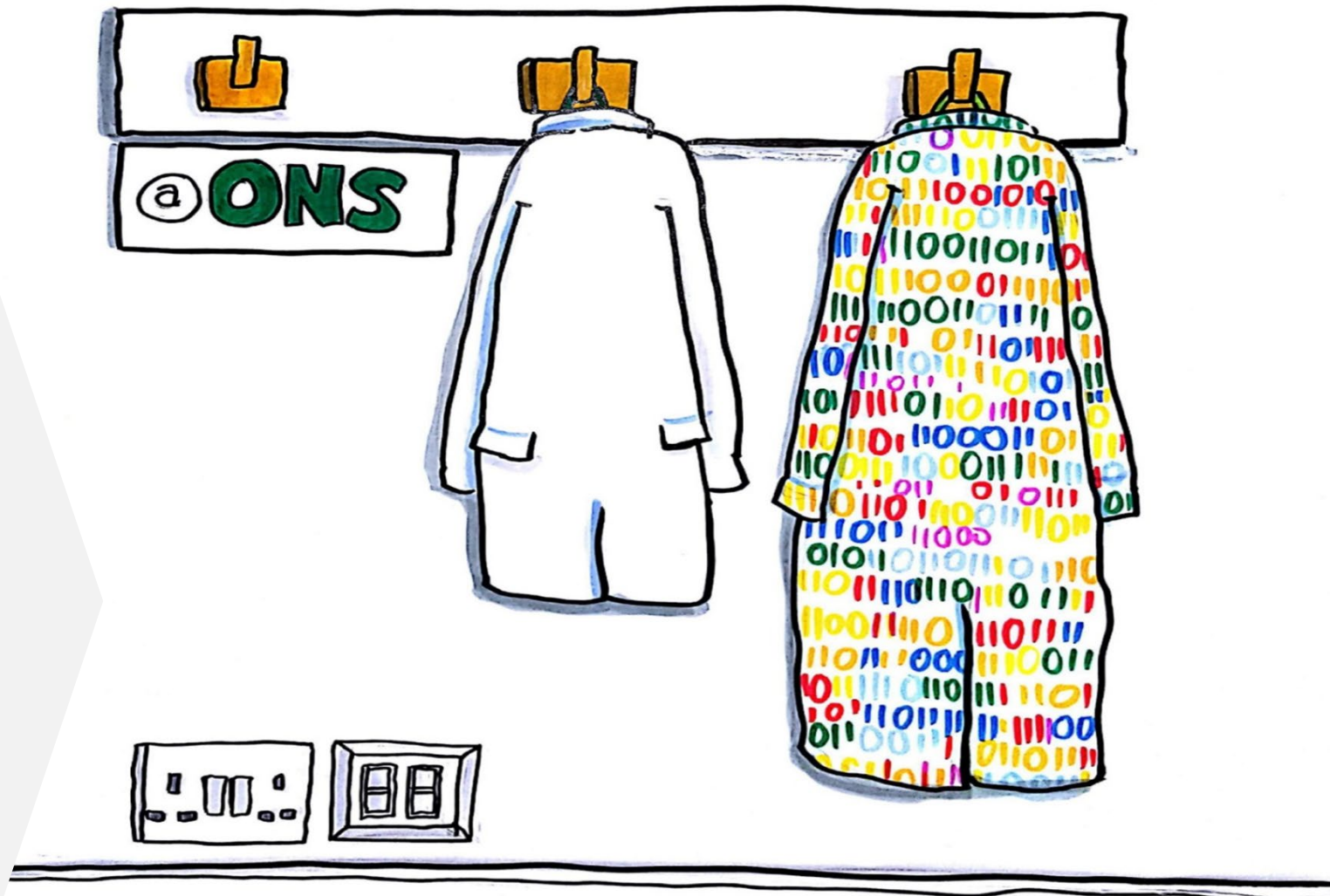


Image credits: ONS stock, Dave Johnson; Louisa Nolan; GDS p2; DFID p17; drawnalism p20