



# The All-Party Parliamentary **University Group**

**5 December 2018**



## **Agenda**

**The subject of the meeting of the All-Party Parliamentary University Group is Big Data in Higher Education**

- |                |  |
|----------------|--|
| <b>8:30am</b>  | <b>Lord Phillip Norton</b> , Co-Chair, All-Party Parliamentary University Group    |
| <b>c8:35am</b> | <b>Dr Paul Feldman</b> , Chief Executive, JISC                                     |
| <b>c8:45am</b> | <b>Richard Puttock</b> , Head of Data, Foresight and Analysis, Office for Students |
| <b>c8:55am</b> | <b>Duncan Ross</b> , Data and Analytics Director, Times Higher Education           |
| <b>c9:05am</b> | Questions, comments, and discussion with university leaders, MPs and peers         |
| <b>10:00am</b> | Speaker meeting concludes.   |

## Speaker Biographies

### Dr Paul Feldman



Paul has been chief executive of Jisc since October 2015.

In January 2016, he was listed in The Sunday Times list of Britain's 500 most influential individuals in the area of engineering and technology.

Before joining Jisc Paul was most recently an executive partner at Gartner UK, a technology research and advisory firm.

He has spent over 20 years in financial services at Nationwide Building Society, Barclays Bank and First Data EMEA, both in IT and business roles.

Paul has also worked in knowledge-based IT companies including Thomson Reuters Legal UK and the Intellectual Property Office.

## **Richard Puttock**

Richard Puttock is head of the Office for Students' data, foresight and analysis directorate. He oversees the OfS' data and analysis work, including management information, data assurance, and sector financial analysis.

Prior to taking up his role at the OfS Richard worked in higher education data with HEFCE for 24 years.



Richard has played a key role in most aspects of higher education (HE) data in his career being actively involved in definition of all of the HESA data streams, many HEFCE data collections and the National Student Survey. Richard is a member of the HE data landscape steering group.

## **Duncan Ross**



Duncan is currently Data and Analytics Director at the Times Higher Education.

Duncan is also Data and Analytics Director for TES Global, helping to create new and innovative data products across the HE and education sectors.

He is listed in the top 50 data leaders in the UK by Information Age and as one of DataIQ's Big Data 100.

Prior to joining TES Global Duncan was Director of Data Science at Teradata, with a remit across all industries and the entire international region (EMEA and APAC), and worked at Experian UK as their Data Director, where he developed their strategic Data Council.

Duncan was part of the team that founded the Society of Data Miners.

He holds an MA in Natural Sciences from the University of Cambridge.

# **Briefing: Big Data and Higher Education**

**Prepared for members of the All-Party Parliamentary University Group**

This is not an official publication of the House of Commons or the House of Lords. It has not been approved by either house or its committees. All-Party Parliamentary Groups are informal groups of members of both houses with a common interest in particular issues.

This briefing document has been produced by Universities UK which provides the secretariat for the University APPG.

## **Introduction**

This meeting of the All-Party Parliamentary University Group will explore how big data is being used for innovation in the higher education sector, with contributions from speakers with a range of views and experiences.

The expansion of datasets that are available on student outcomes, funding and provision have been celebrated by universities, as it allows for in depth investigations in to the impact that universities have and draws attention to possibilities to expand outreach and delivery. The new data landscape is far more complex than before, therefore while new opportunities are abound it has also created difficulties for some institutions in developing the infrastructure to harness this data effectively.

This guide will give further information about the ways in which big data is being used to drive the sector forward, as well as looking ahead as to how data may be used in the future.

## **The Office for Students and Regulation**

The Office for Students (OfS) was established by the Higher Education and Research Act 2017 and came into operation in April 2018. It consolidates the higher education regulatory environment into one organisation. Whereas the Higher Education Funding Council for England (HEFCE) operated as a funding organisation, paid for by government grants, the OfS is funded by registered higher education providers. Providers are required to meet several initial and ongoing registration conditions. Additional specific conditions can be applied to a provider dependent on calculated risk.

The OfS aims to be a student focused and evidence-based regulator, using data and information to underpin its regulatory approach. The OfS defines data as structured provider data returns, annual survey data, data from other bodies such as UCAS and social media or web analytics data. At the time of publication, the OfS' data strategy has not been published, but is expected in Winter 2018. This strategy will expand on the regulatory framework and specify its data requirements for providers.

The OfS' risk-based approach to regulation is a significant change for the sector. This approach will be guided by data, signalling which providers require auditing, the designation of specific conditions and emerging risks for the higher education sector. The OfS does not set numerical targets for providers to meet, but instead assesses providers based on data and evidence. Data will also be used to improve equality and



diversity in the sector, inform students' university choices and allow providers to meet the new transparency duty.

The Teaching Excellence and Student Outcomes Framework (TEF) is a mechanism for assessing teaching quality at universities and colleges, and how these institutions ensure good outcomes for their students in terms of employment or further study. The aims of the TEF are to better inform student choice, and to reward excellent teaching. Participation is currently voluntary and around 300 universities and colleges hold a TEF award.

The TEF process is managed by the Office for Students (OfS), and ratings are judged by an independent panel of students, academics and other experts. The panel considers contextual data and a range of metrics to decide on an initial hypothesis. These metrics use data from the following sources:

- NSS data is used to analyse teaching quality, the learning environment and student satisfaction.
- Higher Education Statistics Agency (HESA) and Individualised Learner Record (ILR) data is used to assess student continuation rates.
- Longitudinal Education Outcomes (LEO) data is used to analyse graduate employment and earnings.
- OfS data is used to look at grade inflation and differential degree attainment (note that these are supplementary metrics).

The panel then considers other evidence, including submissions from the institution, before deciding on a final rating.

## **UK Research and Innovation**

UK Research and Innovation was established by the Higher Education and Research Act 2017 and came into operation in April 2018. The organisation brings together the seven research councils together with Innovate UK and Research England, which performs some of the functions formerly undertaken by the Higher Education Funding Council for England, including lead responsibility for the Research Excellence Framework (REF).

The REF is a process of research assessment, designed to secure the continuation of a world-class, dynamic and responsive research base across academic disciplines within UK higher education.

Data is an important way to measure performance, but it must be contextualised and used responsibly. The Forum for Responsible Research Metrics, chaired by Professor Max Lu, Vice-Chancellor at the University of Surrey, supports the responsible use of research metrics in higher education institutions and across the research community

in the UK. The Forum, among other things, gives advice to the higher education funding bodies on quantitative indicators in the REF 2021.

The higher education sector recognises that research data should wherever possible be made openly available for use by others in a manner consistent with relevant legal, ethical, disciplinary and regulatory frameworks and norms, with due regard to the cost involved. To that end, UKRI, Universities UK and the Wellcome Trust are the signatories to the Concordat on Open Research Data.

## **Funding**

The increased use of big data in higher education has come to influence the way that universities are funded. Large-scale quality measurement frameworks such as the REF and TEF are forming part of the basis for funding allocations. The OfS calculate teaching and capital funding by assessing student numbers that are reported by HESES and HEIFES surveys and allocate this in line with previously agreed methods.

The OfS have used data to target funding towards specific areas where it deems the higher education sector is not serving students as effectively as it could. They do this through OfS challenge competitions which provide universities with funding to create innovative solutions to address issues in these areas. To identify the gaps and issues which will benefit from this type of additional funding, they analyse data that they have previously collected and combine this with external expertise and horizon scanning to determine priorities that need to be addressed.

## **Student outcomes**

As higher education has been moving towards becoming a marketised commodity with institutions encouraged to compete against each other to attract students and the funding that they bring, it has become more important for universities to be able to advertise the difference that they make. A crucial aspect of this is specifically looking at the benefit they bring to students in terms of both their satisfaction with their time in higher education and their outcomes post-graduation.

Important for measuring this impact has been the National Student Survey (NSS), which was launched in 2005 and has been managed by the OfS since April 2018 on behalf of the UK funding and regulatory bodies. It gathers opinions from students, mainly final-year undergraduates, about their time in higher education, providing an influential source of public information and giving students the opportunity to shape their course and institution. Nearly 3 million students in total have taken the NSS and more than 70% of final-year students completed the survey in 2018. The data provided mostly gives an insight into students' satisfaction with different aspects of their course

and institution, which can then be used by universities to drive improvements in their provision of education.

The Destinations of Leavers from Higher Education (DLHE) survey collects information on what leavers from higher education programmes are doing six months after qualifying from their course. This can show the benefit that higher education brings to individuals in terms of employability and earning potential, by providing clear statistics of the unemployment rate for graduates as well as the salary levels of those in employment. The data from 2016/17 shows the destinations of 73.9% of all UK and other EU-domiciled graduates in this year, breaking this down by gender, course of study, location of both study and destination and by provider.

The DLHE is set to be replaced by the Graduate Outcomes Survey, which will start this month, organised by HESA. This was launched following a review into the way that data was being collected about graduate outcomes. The new model is designed to capture rich, robust and innovative data about graduates, using a future-proof and efficient methodology. The first round of data is expected to be published in Spring 2020.

A new global employability ranking, designed by HR consultancy Emerging and published exclusively by Times Higher Education, was launched in 2015, with the most recent ranking released in November 2018. To produce this ranking an online survey was completed by 7,000 respondents from 22 countries who cast around 75,000 votes for universities they felt were the best for graduate employability. Together they represent employers that have recruited more than 250,000 young graduates in the past 12 months. This is a very large dataset giving an insight into the opinions of employers regarding the value of degrees and therefore another measure by which to rank universities.

Recently there has been a movement away from measuring graduate employability with such a short-term focus, instead considering the outcomes of graduates after a more significant period of time. Longitudinal Education Outcomes (LEO) data enables us to know how much UK graduates of different courses at different universities are earning now, either one, three or five years since graduating. It does this by linking up tax, benefits, and student loans data. The data can be broken down by graduate characteristics including gender, ethnicity, region (at application date), age (when commencing study) and prior school attainment. This can be very useful for prospective students to have a better picture of the labour market returns likely to result from different institution and course choices.

While all of these are useful resources for comparing the work of universities they do not help us identify the universities with the best or most effective teaching, nor of the

‘value added’ by a university degree. Graduate outcomes are not necessarily a performance indicator of a university as so many external factors are also involved. Equally, a graduate’s salary should not be the only marker of their relative ‘success’.

## **Measuring impact**

Data is essential for measuring the impact that universities are having in the UK. It can be used to observe how many doctors they are training each year and how many jobs they provide for the economy, as well as contributions for the public such as how many public lectures are held each year and how many students volunteer in their local community.

Measuring this impact serves as a very persuasive argument for the benefits that universities provide. A major area within this has been the financial gains that international students bring with them to the UK. Research undertaken by Universities UK has shown that in 2014/15 international students spent £5.4 billion off-campus on goods and services, which is a massive boost to businesses all over Britain. Without the use of data, it would be impossible to show that this occurred and recent advances in the methodology of collecting this information have allowed for more targeted research into the specific benefits of different aspects of universities.

Times Higher Education has started collecting data for a new ranking that will be the first to measure global universities’ success in delivering the United Nations’ Sustainable Development Goals (SDGs). The first edition of the ranking will include metrics based on 11 SDGs, but the long-term goal is to measure performance against all 17 goals. The data collected from universities will be combined with data from Elsevier to produce an overall ranking of universities as well as separate rankings of institutions that are best achieving the individual SDGs. This demonstrates the many innovative ways that data is being used to monitor the impact that universities are having.

## **Outreach and Access**

Analytics can be a powerful way to identify students who are struggling, and when linking this with demographic data, it can provide insights to particular issues faced by certain groups.

The OfS collect and use data and other evidence to help the higher education sector understand the issues surrounding access and participation. Data has been collected to measure the proportion of students from different backgrounds that are entering higher education. One aspect of this is the participation of local areas (POLAR) classification groups across the UK, which looks at how likely young people are

to participate in higher education across the UK and shows how this varies by area. They also hold data which reports the number of students in higher education by gender, age, disability, ethnicity.

Data has also been used to assess the achievements of students from different demographics during their time in higher education. This looks at students by certain personal characteristics, but also based on past-achievements and qualifications held to determine how likely they are to achieve successful outcomes at university.

An area that has drawn more attention recently regards non-continuation rates and transfers, with the OfS tracking movements of students out of and between higher education institutions. They have found non-continuation overall has increased from 7.4% for entrants in 2014-15 to 7.6% for entrants in 2015-16. This increase is consistent for male and female students as well as for young and mature students.

All of this data can be used by universities and regulators to expand opportunities for young people from disadvantaged backgrounds to enter and succeed in higher education. Not only can it be used to ascertain those groups which are underrepresented in the sector generally, as well as at certain institutions or in certain courses, but can also isolate the most effective methods for reaching these groups, or techniques for providing them support before, during and after their time at university.

## **Next Steps**

More and more data is being generated, across both the public and private sector. This growth presents opportunities and challenges. New methods and approaches are required to make best use of data, and universities play a key role in the development of those methods. The Alan Turing Institute, created by five founding universities and the Engineering and Physical Sciences Research Council (EPSRC) is at the cutting edge of data science. The Institute, which has been joined by a further eight universities in 2018, is well placed to draw on the strengths of the UK's research to take the next steps in data science and artificial intelligence.

Universities also collect significant amounts of data charting students' footprints through their studies and extra-curricular lives; these datasets are rich and growing. Technological and methodological leaps mean that collection and analysis is easier than ever. However, this data resource has often been underused. A 2015 study of 53 institutions by the UK Heads of e-learning Forum found that nearly half of the institutions surveyed had not introduced learning analytics at the time and variable awareness and understanding of the benefits across disciplines.

Learning analytics provide a set of powerful tools to inform and support learners. They enable institutions and individuals to better understand and predict personal learning needs and performance. The UK has one of the highest levels of expenditure per student among OECD countries. Advanced analytics can allow institutions to move beyond narrow performance targets to better understand impact. For example, predictive learning analytics can be used to inform impact evaluations, via outcomes data, that allows institutions to focus resources on effectively support students.

As learning analytics software becomes widely used across the sector, we can expect to see increased demand among the student body for dashboards and analytics specifically targeted at them, rather than designed predominantly for use by tutors. Some UK institutions are already exploring student-focused learning analytics applications, such as Nottingham Trent University's student dashboard. Jisc's learning analytics model includes a student app, which will give students the ability to compare their progress with others through an activity feed, their performance history and a function through which students can set learning targets.

The growth of the open data agenda is likely to continue, as universities make their large datasets available for others to access, increasing opportunities for collaboration. HESA are driving this change having created an open data strategy that will run until 2021, with plans to release data on staff in HE and finances of HE next year, and the destinations of leavers from HE in 2020. We expect this trend for open data to continue and facilitate the more collaborative use of data in the future.