

assessment of science and scientists? As James Wilsdon, lead author of the UK report, writes in a World View on page 129, the one certainty in this debate is that the lure of metrics will only increase. Scientists should not stick their heads in the sand and pretend that the issue will go away. Rather, they should engage with metrics and work to improve the evidence base for them.

British universities now track the output of their academic staff using systems to gather details about their funding and types of output — patents, papers, citations and research grants — and to analyse institutional strengths for comparison with rival universities.

A sophisticated infrastructure has sprung up to support this activity. But it is patchy and inconsistent, with university managers often hopping between various approaches. Some, for example, have built their own internal research-information systems, and others rely on online databases of researcher outputs collected by funding agencies. There are non-profit systems that use public information, and commercially owned databases of bibliometric citations. A host of commercial benchmarking services can analyse the information. These analytical services are becoming increasingly sophisticated. They feature many different ways to group citation metrics, to cover collections of papers by individual, department, institution or journal, and to benchmark them against similar groups.

The problem is that most of these metrics tools lack transparency. At the heart of the system, databases of academic outputs and citations are not publicly accessible or auditable. And the indicators built on top of these databases can also be black boxes: the UK report notes that there are no fewer than ten major global rankings of universities, for example. Some use poorly explained scores and arbitrary weightings

to underpin their league tables, and as the report says, they “assume degrees of objectivity, authority and precision that are not yet possible to achieve in practice”. To some extent, metrics are used and quoted simply because other universities use them — the supply of league tables creates its own demand.

Such opacity can lead to distrust, negating the very advantage of metrics over qualitative assessment as objective, open measures of research performance. It is essential, therefore, that universities are open about the metrics that they build and use.

**“It is essential that universities are open about the metrics that they build and use.”**

Transparency is one of the hallmarks of ‘responsible metrics’, a term introduced by the report that covers principles such as using robust data and applying diverse indicators that account for variation by field and for multiple research types. Other principles include being humble about the limits of quantitative evaluation — which the report notes should support, rather than replace, expert assessment — and recognizing that indicators must change over time.

Although it seems legitimate to use a range of metrics to analyse research performance, their use as managerial targets can leave academics feeling ‘painted by numbers’ — requiring them to change their behaviour to meet often-arbitrary goals. Institutions should therefore publicly state their principles to research managers and explain why they are using particular indicators as a management tool, as the report recommends. Perhaps the most important aspect to recognize about metrics is that they can make judgements more objective — but they can also objectify those being judged. ■

## Cloud cover

*Opposition to storing vast scientific data sets on cloud-computing platforms is weakening.*

Earlier this week, the trade magazine *Computer Weekly* ran a short online news story about a British genomics project buying a commercial system to store its data. Genomics England, which runs the 100,000 Genomes Project, had decided to “reject” the chance to develop its own open-source system, the magazine reported.

In the past, the finer details of IT procurement were not a hot topic for researchers, and so were largely ignored by *Nature*. No longer. Just as a budding journalist cannot hope to flourish these days without a decent working knowledge of the web and multimedia skills, so young (and not-so-young) scientists must increasingly navigate the landscape of large-scale digital-information management.

As a workshop on scientific computing in Portland, Oregon, last month put it: “Computational and data-driven sciences have become the third and fourth pillar of scientific discovery in addition to experimental and theoretical sciences.” In the era of big data, researchers — and journals — simply have to know their HPC (high-performance computing) from their IOPS (input/output operations per second).

‘Big’ barely does justice to the scale of modern scientific data. Mega, giga, tera: all are becoming increasingly familiar — then redundant — terms as the sheer colossus of research data continues to build. The Large Hadron Collider at CERN, Europe’s particle-physics laboratory near Geneva, Switzerland, can generate some 25 million gigabytes of data each year — around ten times the estimated storage capacity of the human memory. Just where are we going to put it all?

A new destination has emerged in recent years: stick it in the cloud, the pervasive web-based services that will, for a fee, take your files off your hands. Late last month, the Broad Institute of MIT and Harvard, a biomedical and genomic research centre in Cambridge,

Massachusetts, announced a partnership with Google Genomics to use its cloud-computing platform to store, analyse and share data. Other clouds are available, and scientists have hooked up with many of them.

In a Comment piece on page 149, several senior scientists call for this trend to accelerate. Major funding agencies, they say, should pay to place large biological data sets on the servers of the most popular cloud services — Google, Amazon and Microsoft among them. Authorized scientists would then be able to tap easily and relatively cheaply into this “global commons”. The US National Institutes of Health (NIH) has cleared the way for such a move: earlier this year it lifted its 2007 ban on using cloud computing to keep and work with its own genetic database.

The NIH had been anxious about the possible threat to the privacy of those who had submitted samples. Such concerns are even more acute in Europe, where the European Commission is already engaged in an ambitious effort to crack down on how personal information is used online. (Scientists have flagged concerns that proposed new data-protection regulations could inadvertently damage clinical research; see *Nature* 522, 391–392; 2015.) So it is reassuring that the commission has pledged to increase the access to scientific data through a continent-wide cloud-computing platform.

As we report in a News story on page 136, plans for one possible model for the European Open Science Cloud are gaining momentum, following a meeting in Geneva two weeks ago. Supporters of the project say that it would reassure academics who are reluctant to use commercial cloud services for security reasons or for fear of being tied to a particular provider. Some of the millions of gigabytes produced by CERN have already gone into a prototype system called the Helix Nebula Marketplace, involving commercial IT providers, and the lab is among those pushing for the idea to be scaled up. As a striking graph published in the Comment illustrates, the number

**➔ NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xhunqv](http://go.nature.com/xhunqv)

of geneticists using cloud-based services is rising rapidly. Astronomers and other researchers are doing the same. At the very least, almost all researchers should explore the options. For more, watch this space. ■