Berlin Open Access Conference
19th November 2013
David Willetts, Minister for Universities and Science

It is a great pleasure to be here today, and I hope that this tenth anniversary of the seminal Berlin Declaration will provide the impetus to push forward a truly global initiative on the important goal of open access.

I would like to thank the Max Planck Society for convening another of these highly useful international summits – a chance to see how one's country's efforts compare with others, and to discuss ways of overcoming common practical problems. There is no doubt that the Max Planck institutes have become synonymous with excellent research and world-leading scientists. With more than 15,000 articles a year in internationally renowned scientific journals, and many of them highly cited, you are leading from the front on the mission to open access to research.

It is a privilege, too, to be here at the Berlin-Brandenburg Academy of Sciences and Humanities – an institution with 300 years of scientific excellence behind it. Both our hosts embody the multidisciplinarity that is at the heart of modern research. None of the complex challenges that we face in the world today – climate change, terrorism, an ageing population – can be solved through the lens of one discipline alone. And therefore institutions such as yours are very much the future.

Germany has a long and impressive pedigree when it comes to science. At the beginning of the 19th Century Wilhelm Humboldt laid the foundations for the modern research university, bringing together research and teaching. Germany also created the first industrial R&D lab. And in the early 20th Century your impressive network of public research institutes - Fraunhofer, Helmholtz and of course Max Planck – created a vital model for research around the world.

Some of the most important discoveries in science in the early 20th Century were made in Germany, including of course Max Planck and his quantum theory which forms the basis of modern physics and is now revolutionising many other areas.

And what has Britain's contribution been? In his book the Gifts of Athena the economic historian, Joel Mokyr, argues that one of the reasons why we had the industrial revolution in Britain was that we had the means to communicate new ideas effectively. Our rich network of learned

societies, publications, and a lively literary life meant that information spread, and new technical advances were more rapidly absorbed across British society than any other country. That carries on to this day. And the underlying argument applies to open access to research today in just the same way.

The Berlin Declaration acknowledged that in terms of both technology and popular culture, the game had moved on. It pointed out that "the Internet has fundamentally changed the practical and economic realities of distributing scientific knowledge and cultural heritage". And it established the important principle that scientific communication could no longer happen in an enclosed academic universe.

This is open access in a nutshell. The problem with the old system is that there are insiders and outsiders. The insiders are of course the academic community, who produce research in an exclusive space. Meanwhile, with their noses pressed to the window we have the outsiders – who have paid for the research with their taxes but nonetheless cannot afford to access it. The outsiders include independent researchers and authors, not affiliated to a university or research institute but doing valuable work nonetheless, small business owners who want to keep abreast of developments in their area, and even students at less well resourced universities.

I had my own experience of being an outsider when I was writing my book, The Pinch, on fairness between the generations. It was very frustrating to track down an article drawing on publicly funded research and then find it hidden behind a pay wall. That meant it was freely accessible to a professional in an academic institution, but not to me as an independent writer. That creates a barrier between the academic community – the insiders - and the rest of us, which is deeply unhealthy.

I think that most of us in this room want to get past that polarised world. But the real challenges of doing so are embodied in the final sentence of the Berlin Declaration.

"Our organizations aim to find solutions that support further development of the existing legal and financial frameworks in order to facilitate optimal use and access."

I convened a meeting of the G8 science ministers in London in June and I made sure that open access was on the agenda. It was these challenges that everyone was debating.

And it is these challenges that I wish to tackle today. I want to update you on exactly where we are on open access in Britain, and to explore the implications of the choices that are being made internationally about whether to opt for Gold or Green.

And I want to look at a parallel issue of equal concern – data linking, and access to the data underpinning research papers. If science is to be truly replicable there are some serious problems that need to be addressed here.

But let me start with where we are on open access in the UK.

We made a firm commitment to public access to taxpayer-funded research two years ago, in December 2011, in our Innovation and Research Strategy. To take this forward we formed a national working group to find practical ways to broaden access, led by Dame Janet Finch. The group brought together all the key players, including the research community, funding bodies, publishers and information providers.

Dame Janet Finch published an excellent and bold report in June last year. Our Government endorsed all of her recommendations with the exception of one point on VAT on e-journals which did not concur with European rules.

Research Councils UK published its guidance on how it would implement this new policy in April. This confirmed that they would accommodate both the Gold and Green open access models, but that the clear preference was for immediate or Gold open access with the maximum opportunity for search, access and re-use.

To accelerate take-up of the Gold open access route, the Research Councils have provided block funding to the UK's universities and research institutions to allow them to fund their researchers' Article Publication Charges. They will invest £17 million in over 100 Publication Funds in 2013-14 and a further £20 million in 2014-15. Universities for their part are investing significant effort in their internal systems and processes to enable these funds to be effectively used.

I want to explain the logic behind that important decision to go for Gold so explicitly in some detail.

I understand that on the surface at least going for Green looks seductively simple. Having decided to move towards open access, it might seem simple to say that the cheapest way of delivering this is to require articles

to be made available outside paywalls more quickly. But the major problem with this is that it fails to protect either the timeliness or assured quality of what the reader is seeing or the publisher's ability to deliver peer review to achieve that quality in the first place. And any country that values high quality and reliable science underestimates the worth of the publishing industry at their peril.

We are very proud of our publishing industry in Britain – which includes learned societies as well as major commercial players. But this is a global success story. Macmillan, for instance, was founded in the UK in 1843, but is now part of the Holtzbrinck Publishing Group, a family-owned company with headquarters here in Germany. In total the academic publishing industry print some two million peer-reviewed articles per year in about 25,000 journals.

However, academic publishing also adds value to all our countries in a much more fundamental way, because peer review and publication is such a crucial part of the research process. Academics generously give their time to scrutinise draft articles. And value is added by publishers identifying the academics to conduct peer review, through the editorial function of signalling which research is of the highest worth, and by helping others to find it. I do not scan the cacophony of opinion on the web to try to work out what is happening in the world: my laptop is set to open at the BBC news page. This process of winnowing data and research and then ranking them is of great value. Yet of course somehow it has to be paid for – and in a sustainable way.

Gold makes sense because it explicitly recognises that there is a cost and a value to publishing. That is the first advantage.

Plus of course you get the work openly accessible straight away. The hidden cost in green is waiting six months, 12 months, 24 months, whatever it is, until the layperson or SME outside the academic community can access a piece of work that he or she as a taxpayer has already paid for once. That is the second advantage of Gold.

In the UK we have also taken a firm line on embargo length within the Green option. Where the publisher is not offering Gold open access, we will not support articles with an embargo of longer than 6 months for STEM subjects and 12 months for the humanities. However, if the publisher is offering Gold, but the researcher is unable or unwilling to pay the Article Publication Charge, the embargo will extend to 12 months for STEM and 24 months for the humanities. The exception is biomedicine for which we prefer to see publication within six months

under all circumstances, if it can not be published on a Gold Open Access basis.

My personal preference is for a straightforward Gold option – but what I certainly do not believe in is opting just for Green with short embargos. That jeopardises the viability of the publication of academic research.

I was delighted to hear at the weekend that the Dutch have joined the British, the Germans and the Austrians in boldly committing to funding Gold open access. This is tremendous news and will, I hope, encourage other countries to follow suit and commit to Gold.

Europe has hoisted the flag for open access by making it a requirement for all research published under Horizon 2020, when it starts next year. But if the tendency is to go for Green instead of Gold in many countries as the superficially easy option, I think collectively we will come to regret it. Especially if the EU maintains its position of allowing only 6 or 12 months embargo periods even when there is no funding to pay for Gold.

We recognise that this is not going to be an overnight process.
RCUK has made clear that the transition will probably take around five years, and they will review impact next year. However, we are already making good progress.

The Publishers Association has recently reported that in only a year 70% of journals now publish Gold or include a Gold open access option. Of these, 82% allow the author to choose an accommodating CC-BY licence. 96% of journals have an embargo period of 24 months or less, and 64% of journals have an embargo period of 12 months or less.

However, having taken the crucial strategic decisions let me tell you about some of the subordinate problems that we are wrestling with. These are the tricky questions that I'm sure are being debated in the corridors of this conference and they have no easy answers.

The first is the hybrid publication issue. Unlike in Germany the UK allows hybrid models, because we believe this encourages publishers to move much more quickly, and accelerates the move to open access. It also allows researchers to continue to use the branded journals they value. The downside is that this opens up the potential accusation that journals might be 'double-dipping'.

If you are a research intensive university and you are complying with the Gold open access policy enthusiastically you may find that very quickly you are racking up Article Publication Charges in the tens of thousand pounds or even in the millions. In the UK you will receive a block grant to help cover this. But of course at the same time journals are publishing research from all around the world, much of which is not on a Gold basis, and the only way of keeping up to date with that research is for your institution to subscribe to the journals. It is easy to imagine the frustration of the university administrators who find that they are paying a hefty charge to the publisher in Article Publication Charges, and a further hefty fee for subscriptions. And it is easy to see why publishers offering global discounts retrospectively – rather than looking at individual institutions' contributions – might not assuage that anxiety.

However, publishers may find that the market takes its own action if they ignore these concerns. Individual researchers are free to decide how they spend their block grant, and that could include declining to pay article charges to specific hybrid Gold journals that they feel are guilty of "double-dipping".

The second related question is the pricing of Article Publication Charges. In the UK we haven't intervened in the prices charged, preferring to leave it to the market. In contrast Germany has introduced a cap. I anticipate that we will see increased competition in the research publications market which will drive costs down. Researchers will decide when opting for Gold open access which publication they are willing to pay to be in, and those publishers who are seen to be providing value for money will be the ones who flourish in this new market. However, there is no doubt that this is tricky terrain – and a source of anxiety for many institutions that we should not ignore.

I am delighted that an international consortium of funders, including the Max Planck, has commissioned research by Bjork and Solomon to analyse the developing market for open access Article Publication Charges. This will be a valuable input to the RCUK review late next year, by which time we are expecting publishers to have developed more imaginative and transparent ways of reassuring research institutions that their pricing is competitive and based on sustainable business models. It will be up to them to prove that the UK has been right to adopt a policy of allowing hybrids rather than regulating these costs.

Finally, let me talk about our next moves on open access.

First, following on from the constructive discussions held at the G8 science meeting in June, it has been agreed that there should be a follow-up meeting to talk about best practice and to debate shared concerns. This will be held at the Royal Society on March 20th, and I have high hopes that we can drive forward a more unified international approach.

Secondly, we understand that open access can seem complicated. For this reason we have published a Decision Tree, which sets out clearly for researchers and publishers what choices they need to make on open access. I would encourage other countries and the EU to adopt it.

But crucially, thirdly, we are also trying to make the world of research easier for the general public, and smaller businesses in particular, to navigate. We are developing a Gateway to Research, which will allow people to search for the latest research in areas that interest them, as well as linking them up with relevant research teams who might be able to help them investigate or commercialise ideas. You can visit a beta version of this site online now. I hope to launch RCUK's fully developed version shortly.

And there is a fourth major issue that we are focusing on. Beyond open access we need to think about open data.

Last month the Economist set out a challenge to modern science that the fundamental tenet of scientific replicability was not being delivered as we had assumed. This is a concern we must address head on. In an era of massive data volumes, reproducibility and self-correction remain fundamental principles of modern science but are more and more difficult to maintain.

As the Royal Society set out in a report last year, a great deal of data has become detached from the published conclusions that depend upon it. This means that the two vital complementary components of the scientific endeavour - the idea and the evidence - are too frequently separated.

One of the advantages of the Gold model of Open Access is that it enables CC BY – the most useful licence in terms of developing and disseminating knowledge, which lets others search, distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. And the ability as part of gold to have proper access to the data behind the research is another real advantage.

There are real challenges here. For one, there is a lack of coordination and standardisation of data across institutions within particular disciplines. Addressing this will require leadership from learned societies, but will also require institutions to think about the skill mix of both researchers and their librarians. What will the library of the future look like?

There is also the problem with the rather unappealing name "link rot". I don't know how many of you have come across link rot, which sounds like a nasty foot disease. In fact it's the name that refers to the experience of clicking on a URL reference in a paper and reading only an infuriating message saying "page not found". And if that is the ultimate empirical basis for claims in an academic paper, if more and more of the references cannot be found, then that is a significant erosion of our academic research base. There has been some work done on link rot in the Journal of the American Medical Association, the New England Journal of Medicine, and Science – three very prestigious titles which shows that there is 4 per cent link rot after 3 months, 10 per cent after 15 months and 13 per cent after 22 months.

As more and more scholarly material is made available on the web - including via different forms of open access publications - 'link rot' is a problem that is likely to grow. The British Library believes that the most satisfactory way of dealing with this issue would be for citations to point not to live internet pages but archived versions of the same page – which would then be much more reliably preserved.

The first step to realising this should include a consultative process with the research community to determine what such referencing system should look like and how scholarly use of web citation should operate. I therefore intend to invite the British Library to the next meeting of our UK Research Sector Transparency Board which I chair to get this process moving.

At the G8 science meeting we agreed a set of principles on open scientific research data. This stated that the default position of the results of publicly funded scientific research data should be one of open data, subject to concerns of privacy, security and commerciality. This is a strong message and I am delighted that the G8 is united behind this principle.

In the UK the Royal Society has agreed to convene an Open Science Data Forum in early 2014 to seek practical ways forward on the issue of data

standards, and other challenges such as skills, training, and career development for researchers and data specialists in the sector. Its aim will be to engage the parties who will be key to the implementation of an open data regime in research, including representatives of universities, the Research Councils, and other experts.

The prize is huge. CanSAR was launched by British scientists recently. This is the world's largest cancer database, which will use computer programmes similar to those that forecast the weather to process millions of experimental results in seconds. It has the potential to revolutionise the search for cancer cures. In total, 1.7 billion experimental results will be available, free of charge, to researchers all over the world. The database uses artificial intelligence to help scientists understand work in other fields and also find genes that affect cancer, an expensive and time-consuming process. Research that previously took months could now be done in minutes.

Open access and open data are challenges that the whole world must embrace. We all understand the goal. Now we need to work together to make it a reality.