

IMPERIAL COLLEGE LONDON

Response to Cabinet Office consultation on 'Making Open Data Real'

1. The government is rightly finding ways to improve public service, encourage competition and promote economic growth by giving access and free use of publicly funded datasets for these purposes. The college is concerned, though, that the application of these principles to universities (which are private charitable bodies, though partly publicly funded) will have unintended consequences which will work in the very opposite direction to the policy intended.

The consultation asks: How do we get the right balance in relation to the range of organisations (providers of public services) our policy proposals apply to? What threshold would be appropriate to determine the range of public services in scope and what key criteria should inform this?

2. Whether higher education institutions (HEIs) are included in scope will be a crucial decision. The generally accepted understanding of a public service would not normally include post-compulsory education or research. Universities are autonomous charitable bodies not subject to direction by a government minister (indeed such intervention is explicitly illegal). With the forthcoming change in the funding of undergraduate students, grant payable by the Higher Education Funding Council for England will form only a minor part of the cost of higher education for most universities. With the majority of income coming from or on behalf of students as individuals, the government have already strengthened the power that consumer choice will have on competition between institutions.
3. 'Dataset', as defined in the consultation, "will typically have been collected as a by-product of delivery". Universities already have published, or are about to publish large sets of data about their performance. The National Student Survey (NSS) asks final year students about their experience across a number of different aspects of their education. The data are published in a single place for all universities. Likewise, Research Assessment Exercises every few years grade the quality of research in almost all departments of every university. These results are also published for all universities in a single place. Shortly, each university will publish a Key Information Set of data against a common standard intended to inform potential applicants about the attributes of each course such as learning and assessment requirements, staff contact time, the NSS score and the likely employment and salary prospects after graduation.
4. HEIs are subject to FoI requests, despite being private bodies. We would be concerned if an extension to FoI as envisaged in the consultation resulted in further demands for datasets when, in our view, the accountability, competition, choice and quality enhancement policy imperatives are already being met through existing public datasets. The consultation asks what threshold would be appropriate to determine the range of public services in scope. **We do not think that higher education and research is a public service in the normally understood meaning of that term**

and so universities should not be in scope. A helpful additional criterion might be the proportion of funding a private body receives from the government through grant or contract. If this were less than 50%, we would argue that the body in question was out of scope in the same way that, for example, universities are in respect of EU procurement regulations.

An enhanced right to data?

5. We are extremely concerned that core research and teaching datasets in universities, not 'by-products of service delivery', would become subject to open access and be available for commercial use. Notwithstanding the definition of 'dataset' which excludes information in non-government bodies about aspects unrelated to public service, because universities are subject to FoI, the risk is that such data will be brought into scope unless explicitly excluded. One of the aims of open access is to promote economic growth, yet the very opposite is likely to happen in some sectors if this policy is applied without modification. Much university research, whether funded publicly or privately, will result in the creation of large datasets. All the outcomes of research are published, usually in peer-reviewed journals, and publicly available (increasingly in open-access journals). In many cases, though, the datasets themselves will form the basis of a commercially exploitable opportunity. It is essential that these datasets are protected so that they can be commercialised by the university (or its agent) *ie* universities should be encouraged to continue to follow the very path that government is advocating in this consultation.
6. These core datasets should not be taken and exploited free of charge by a third party who had no interest or risk in their creation. Crucially, companies sponsoring university research which generates data may cease to do so if they cannot protect the products of the research. Such companies would be likely to take their funding overseas to jurisdictions which allowed them to protect their assets. This university alone undertook £M130 of research in 2010-11 from private organisations (companies and charities). The loss of such income and the threat that research data might be freely open for exploitation by others would precipitate a loss of many and key staff with consequent damage not just to this university but to all similar ones in England. We have similar concerns about datasets used in teaching, produced at considerable cost, which would then be available without charge for use by a competitor. **Core research and teaching datasets need to be excluded from these proposals as of right, not on a case by case basis using specific exemptions.**

Resource implications other than cost

7. The consultation makes hardly any mention of the overheads that free access brings. These are not simply the cost of making them publicly available but the subsequent cost of supporting them when in the public domain. As well as access for public benefit, these proposals would give access for private benefit. Many single issue pressure groups and journalists on 'fishing trips', for example, would seize the opportunity of accessing specific data, perhaps when not complete. The cost to the university would be in explaining the context, justifying the data, and responding to queries which can only divert academics and others from their productive core work. Although we note that ministerial decision making is exempted, no such exemption is given in FoI (in England) to other kinds of work in progress. The presumption in the consultation is for early release over improved quality of data. The risk (which may

be considerable) is that users make the wrong decisions to their own detriment because of poor data.

Possible unintended consequences

8. A similar risk applies to use of raw data. For example, the Dr Foster company (quoted in the consultation) apply their own knowledge and expertise of the health sector to transform raw mortality data into something useful by the general public. Someone looking at the raw data alone would not necessarily discover which the safest hospitals were for their particular condition and could make a poor choice of hospital to attend. Who would be liable in these circumstances? Would a simple prior disclaimer of liability be sufficient? The whole question of legal liability is not addressed in the consultation.

IMPERIAL COLLEGE LONDON

25 Oct 2011