# Annex A
# Consultation response form for the Research Excellence Framework

1. Respondents should complete the form below.

2. Responses should be e-mailed to refconsultation@hefce.ac.uk by **Thursday 14 February 2008**. HEIs in Northern Ireland should send a copy of their response to research.branch@delni.gov.uk

3. Institutions wishing to express an interest in taking part in the pilot of the bibliometrics indicator should e-mail their details to refconsultation@hefce.ac.uk by Thursday 31 January 2008.

4. We will publish an analysis of responses to the consultation. Additionally, all responses may be disclosed on request, under the terms of the Freedom of Information Act. The Act gives a public right of access to any information held by a public authority, in this case HEFCE. This includes information provided in response to a consultation. We have a responsibility to decide whether any responses, including information about your identity, should be made public or treated as confidential. We can refuse to disclose information only in exceptional circumstances. This means responses to this consultation are unlikely to be treated as confidential except in very particular circumstances. Further information about the Act is available at www.informationcommissioner.gov.uk.

## Respondent's details

**Are you responding:** (Delete one)

- On behalf of an organisation

| | |
|---|---|
| **Name of responding organisation/individual** | ……British Classification Society…… |
| **Contact name** | ……Prof. Fionn Murtagh…… |
| **Position within organisation (if applicable)** | .…President…….…… |
| **Contact telephone number** | ……01784 443429…… |
| **Contact e-mail address** | ……fmurtagh at acm.org…… |

## Consultation questions

**(Boxes for responses can be expanded to the desired length.)**

**Consultation question 1a**: Do you endorse our proposals for defining the broad group of science-based disciplines, and for dividing this into six main subject groups, in the context of our new approach to assessment and funding?

The division between science, engineering, health, and medicine (called science-based disciplines) versus humanities, social science, mathematics and statistics (called non science-based disciplines) will require somewhat varied strategies by researchers in the interests of inter-, multi-and cross-disciplinarity.

**The British Classification Society represents researchers spanning many of these fields. In the interest of cross-disciplinarity we would like to see flexible, adaptive and constructive harmonization across these different areas as soon as possible.**

**Consultation question 1b**: Are there issues in relation to specific disciplines within this framework that we should consider?

Application disciplines are not necessarily well covered by blbliometric indicators. This applies to engineering and health domains, among others.

Mathematics and statistics are strangely categorized as non science-based disciplines. (See note (1).)

Computer science is categorized as engineering, which is arguably the most appropriate general category, but this does not recognize that it is also (or various subfields are) mathematical; scientific as opposed to engineering; arts-related (e.g. design); and social science related (e.g. psychology, information systems, management science).

**We recommend that the categorization of disciplines and subdisciplines needs to be appraised and updated on an on-going basis, in relation to underlying discipline development and evolution, and national policies and strategy.**

**Consultation question 2a**: Do you agree that bibliometric indicators produced on the basis that we propose can provide a robust quality indicator in the context of our framework?

Bibliometric indicators have become commonplace in most areas of scholarly research, and are used in formal and informal ways in comparative evaluation of researchers, and as performance and innovation benchmarks.

However there are crucial data analysis issues to keep in mind.

Firstly the data will, in many cases, be long-tailed (self-similar) and therefore will be "bursty" on all aggregation scales. The "bursts" here are the infrequent but nonetheless pervasive very large citation values. This precludes any easy recipes for summarization. (See note (7).) Robustness has a precise meaning in statistics, viz. that censoring the data of more extreme and outlying values produces more stable summary statistics. Such censoring brings little benefit with long-tailed distributed data. As an example (see also note (7)), consider how long-tailed data citation distributions have the immediate effect that one highly cited article has far greater bibliometric importance than poorly cited, techncial consolidation work in an archival journal, notwithstanding the latter's possible role in educating future researchers.

Secondly we are dealing with highly multivariate data. One size cannot fit all: one numeric output cannot do justice to the many dimensions in the data; to the fact that the data is of mixed properties – qualitatitive, numeric and textual (among others) on input; the dynamic (time-varying) aspect of the data; and the fact that there are various planned and serendipitous results from analyzing any large and complex data set.

Thirdly fitting models to the data must be local, in other words based on semantic clusters in the data, in order to handle both intrinsic complexity and multiple criterion ranking outputs. (Example: analysis of a subdiscpline.)

**We recommend that both data and data analysis pipelines be taken as integrally related and mutually supporting.**

**Consultation question 2b**: Are there particular issues of significance needing to be resolved that we have not highlighted?

It is unclear if all data will come from Thomson Scientific, given the role of the web, and of services like Google Scholar, ADS, SPIRES, Scopus, CiteSeer, PubMed, PubMed Central, learned society digital libraries like those of the IEEE and ACM, preprint services like arXiv, and others. Issues of data integrity, trust and transparency are considered in [6], and issues of coverage in [3].

**We recommend against one single (commercial, proprietary) provider in favour, instead, of open, publicly accessible data sources.**

**Consultation question 3a**: What are the key issues that we should consider in developing light touch peer review for the non science-based disciplines?

In support of light touch peer review, with the non science-based disciplines, and also in support of the science-based disciplines, a major gain for all research disciplines will be

the data collection. Having high quality data is of inestimable importance. How validation will be carried out, on an ongoing basis, is an important issue for all of us, whether science-based or non science-based. Note that in the RAE, the higher education institutes have the responsibility for data collection and input. In the REF, lines of reporting need to be established.

**We recommend that the quality as well as quantity of data be assessed on a continuing basis. Data is a resource of great value. Data understanding and interpretation cannot be divorced from either end of the chain linking data gathering and selection, at one end, with decision making at the other end.**

The role of book publishing in either science- or non science-based domains has not been addressed adequately. Note that books are rarely open and freely accessible, unlike the vast bulk of the scientific literature (in preprint servers, and in personal and institutional repositories). In [8] it is noted that books and policy reports need to be included in bibliometrics for many areas of social science. Monograph, research-based, book publishing is also relevant for e.g. Computer Science, Mathematics and many other fields.

**Recommendation: open access policies (e.g. [9]) hold for journals and conference publications. There are other outputs of scholarly work that must also be taken into account, including books, reports, data and software.**

**Consultation question 3b**: What are the main options for the form and conduct of this review?

Dissemination of a great deal of scholarly output is via the web. Harvesting this information, including personal, departmental and consortium web sites, is clearly a way to proceed. However the issue of rectifying errors, and assessing redundancy of information (including in published information), need to be considered on an ongoing basis. As examples of redundancy checking, consider citations to preprints, or to web versions of a paper; or conference publication and a subsequent, extended journal paper.

Mechanisms are required for data validation. Redundancy checking should be considered on both syntactic (e.g. version control) and semantic (overlapping content) levels.

**We recommend that all reviews, driven by policy objectives, be based on objective and transparent data analysis, including bibliometric analysis.**

**Consultation question 4**: Is there additional quantitative information that we should use in the assessment and funding framework to capture user value or the quality of applied research, or other key aspects of research excellence? Please be specific in terms of what the information is, what essential element of research it casts light on, how it may be found or collected, and where and how it might be used within the framework.

It is essential to consider applied research outputs including patents and other intellectual property; open source software; company start-ups; university-based commercialization activities; collaborative links with companies including SME and large; industry-supported PhDs.

Measures of esteem, including keynote presentations, programme committee memberships, journal editorial board memberships, leadership roles in learned societies, prizes and awards, fellowships in learned societies, etc., need to be taken into account.

Also to be considered: gender balance; sustainability in terms of age and career stage profiles; profile of early career researchers; intra-discipline coverage and breadth; uptake of sabbaticals; and facility of parenting or carer leave.

Evidently also PhD production, and breadth and depth of PhD training.

**We recommend that there are many other sources of data to be taken into account, expressing the results of research and scholarly work.**

**Consultation question 5**: Are our proposals for the role of expert panels workable within the framework? Are there other key issues on which we might take their advice?

It will be valuable to have available tools for data validation, and cross-correlation of inputs (e.g. redundancy checking mentioned under 3b). Facilitating mapping studies as discussed in [10] is just one example of the gains to be made by use of blbliometric and citation based analysis.

**We recommend that objective, transparent procedures for decision making, based on bibliometric and other quantitative information, represents a major step forward. Expert panels should be informed by such information at all stages of their activities; but informed, not largely supplanted as suggested in the proposals.**

**Consultation question 6**: Are there significant implications for the burden on the sector of implementing our new framework that we have not identified? What more can we do to minimise the burden as we introduce the new arrangements?

It must be recognized that in the current proposals the sector will be left with the responsibility of dividing staff into science and non-science (often quite an arbitrary

division, as noted in the responses above to 1a and 1b), for data collection and for data validation (as noted under 3a, a far greater responsibilty than under the RAE). While the potential of the data is very great, without well established data semantics and easily deployable software toolsets for data verification and cross-correlation (see note (8)), the challenge is very great.

Procedures and outcomes should be seen in a global context in order to benchmark externally as well as internally.

**We recommend that a global research baseline and comparator is needed.**

Continuing use of both metrics and conventional panels in different contexts are a means of ensuring trust in the system, coupled with availing of the benefits of each (e.g. metric-based for smaller funding programmes; panel-based for larger, more dedicated programmes).

**Consultation question 7**: Do you consider that the proposals in this document are likely to have any negative impact on equal opportunities? What issues will we need to pay particular attention to?

It has been noted in consultation question 4 that equal opportunities (and other aspects relating to fairness of treatment) must also be quantitatively assessed on an ongoing basis.

**Consultation question 8**: Do you have any other comments about our proposals, which are not covered by the above questions?

The British Classification Society sees the marvellous potential of having high quality and openly accessible data available to researchers as well as policy makers, and also analysis methods and frameworks (both conceptual, and software toolsets). These will lead to clarity and focus in informing research at all levels. High quality research and R&D outputs and productivity data, understood in the global context, comprise a national resource of high value and of the highest strategic importance.

## Notes and Further Comments

(1) The placing of mathematics and statistics as non-science subjects is unusual. Mathematics and statistics are discussed on p. 13 of [1], where RAE returns in these disciplines in 2001 are compared with STEM subject coverage in Thomson-indexed journals. It is concluded for these disciplines: "… there are data limitations, which are compounded by behavioural factors that produce an unusually high rate of 'uncited' papers."

(2) The data analysis problem from [1], pp. 32-33:

The UK publishes about 100,000 articles a year. … In 2004, those articles had a total of 473,046 authors, not all of whom were in the UK. … There were about 8.5 citations per UK paper on average over the five-year period from 2001 to 2005, or about four million citations to about 500,000 papers.

From other sources, there are about 24,000 peer-reviewed journals, and about 2.5 million articles published in peer-reviewed journals annually.

(3) Issue of data quality, and need to manually check all entries:
[1], p. 32: "there are at least three unique Dr F Guillemot's in UK data".

In [3], the Leiden group, which is currently assessing the use of bibliometrics in conjunction with the 2007 RAE, reports on computer science and how the Thompson Scientific databases, Science Citation Index (SCI) or Web of Science (WoS), cater or otherwise for conference publications or e.g. Springer's LNCS, Lecture Notes in Computer Science, in this area. They [3] conclude (Recommendations, p. xiii) that "the existing Expanded WoS publication database should be further expanded (and updated) with a number of additional important conference proceedings volumes". It is unlikely though that this would raise coverage above about 80 per cent. So "a further expanded WoS database" is called for. Ultimately there is a need for citation analysis to "be carried out – or at least checked – manually" (Recommendations, p. xiii).
The global picture ([3], Recommendations, p. xiv): "… there is as yet no 'standard' methodology available that allows for international comparisons. … further research into the development of such a methodology is necessary."

(4) Pygmalion in the research lab:
[1], p. 35: as early users of biometrics evaluation criteria, the Dutch have become hugely influenced and have reacted accordingly. There has been "an exceptional rise in output and citation share" as a consequence of such "emergent behavioural effects".

Examples of changes of behaviour include the following. Among editors it is known that special issues help in journal citation scores, and thereby too the authors' scores. Journal editors have been known to ask accepted authors to include additional citations from the same journal in the crucial two year period that counts for journal citation ratings. The paper, http://arxiv.org/abs/0801.0386, discusses how cross-citing (possibly

asymmetric) networks are relevant in practice. In some fields US-based authors publish iin, and cite, preferentially US-based publications.

(5) The crucial importance of baselining and weighting data.

[1] makes many helpful points but also is focused strongly on the need to normalize bibliometric data (see pp. 3, 4, 9, 17, 29,30, 33). Data "need to be normalised to account for … the fact that their distribution is skewed". And "it is … necessary to transform them [data with skewed distributions] in some way to arrive at a more intuitive presentation and manageable analysis" (p. 9). And: "Normalisation strategies are a critical part of any metrics-based methodology…" (p. 17). A citation is given to [3], nothwithstanding the critical nature of this article. The crucial role of normalization is pointed to in [5] also.

They data are more than "skewed": they are long-tailed, self-similar.

(6) Message of Stevan Harnad to CPHC (Conference of Proressors and Heads of Computing) email list, 16 Nov. 2007, with subject "Continuous multi-metric research assessment", in reply to Jonathan Adams (Director, Evidence Ltd.).

- "I think you greatly underestimate (1) the power of multivariate (as opposed to univariate) analysis, validation and weighting as well as (2) the power of open access (i.e., online, public, pervasive, continuous, and dynamic) metrics."
- "Think multivariately, dynamically, and openly."
- "Don't think cravenly in terms of how the UK will stack up in terms of existing, unvalidated, univariate metrics. Think in terms of establishing metric standards for the entire world research community in the metric OA era!"

We agree with these points about the multivariate nature of the data, and the need to understand the data in a global context. The issue of Open Access is across many disciplines largely a matter of fact. See [9] for European policy in this regard.

(7) On the issues of long tailed data, consider the following example from [6]:
"Because the impact factor calculation is a mean, it can be badly skewed by a 'blockbuster' paper. […] In a self-analysis of their 2005 impact factor, Nature noted that 89% of their citations came from only 25% of the papers published."

(8) Among the needs expressed by Hafen and Peitsch [11] are: a unique author identification, a UAID analogous to a DOI which uniquely characterizes published material; a global discussion and rating platform; and semantic web support for disseminating and navigating scientific information. The authors are from the bioinformatics and systems biology community.

## References

[1] “The use of bibliometrics to measure research quality in UK higher education institutions”, Research Report, prepared for Universities UK by Evidence Ltd., pp. 40, October 2007.

[2] L Leydesdorff and S Bensman, “Classification and Powerlaws: The logarithmic transformation”, Journal of the American Society for Information Science and Technology, 57, 1470-1486, 2006.

[3] H.F. Moed and M.S. Visser, Developing Bibliometric Indicators of Research Performance in Computer Science: An Exploratory Study”, research report to the Council for Physical Sciences of the Netherlands Organisation for Scientific Research (NWO), CWTS Report 2007-01, Feb. 2007.

[4] Research Excellence Framework, Consultation on the assessment and funding of higher education research post-2008, Nov. 2007/34.

[5] J. Adams, L. Jackson and S. Marshall, Bibliometric analysis of interdisciplinary research, report by Evidence Ltd. to HECFE, Nov. 2007, pp. 35.

[6] M. Rossner, H. Van Epps and E. Hill, “Show me the data”, editorial, JCB – Journal of Cell Biology, 179 (6), 1091-1092, 2007 (http://www.jcb.org/cgi/content/full/179/6/1091)

[7] Z. Corbyn, “Researchers raise questions over validity of REF citations criteria”, Times Higher Education, 10 Jan. 2008, p. 11. Compared to ADS (Astrophysical Data System), WoS (Thomson Scientific World of Science) was 5000 papers short in one case or 18%. In another case, compared to SPIRES (high energy physics), 1600 papers or 37% were lost.

[8] Z. Corbyn, “Social scientists are divided over possible use of citations in the REF”, Times Higher Education, 17 Jan. 2008, p. 7.

[9] Scientific Publication: Policy on Open Access, European Research Advisory Board Final Report, EURAB 06.049, Dec. 2006.

[10] Scoping Study on the Use of Bibliometric Analysis to Measure the Quality of Research in UK Higher Education Institutions, Report to HEFCE by the Centre for Science and Technology Studies, Leiden University, Nov. 2007.

[11] E. Hafen and M.C. Peitsch, “Science in a Web-2.0 environment – draft white paper”, 11 Dec. 2007.