Research evaluation – is it our business? The role of librarians in the brave new world of research evaluation

JIBS User Group - Workshop 29 June 2010 – University of Birmingham

As the tweets (#JIBSre) appear to indicate, this JIBS workshop apparently hit just the right topic at just the right time. People came with all different levels of experience, and being so well-attended the workshop opened up many opportunities for sharing that experience. Whether embedded in your work already or not no-one could have left the day without gaining a wide array of ideas for practical action to take back to institutions. And the venue was the stunning campus of Birmingham University – on a stunning June day.

The workshop's remit was inspired by institutions' interest in procuring a CRIS (Current Research Information System) and the perceived and growing requirement from multiple points of view – both internal and external – for analysis of bibliometric information. Procurement of a CRIS is one solution among many (and there are as many hybrid as PURE solutions) that comprises re-tooling of current internal systems, and alternative uses to those originally intended for institutional repositories. At the centre is the librarian – or is s/he? "Is it our business?" It was one of the day's main aims to contribute ideas to feed into librarians' individual and collective strategies for getting involved in this new area of work; an area that is of crucial significance for institutions as they face the REF in 2014 . Equally librarians find they must step up to calls to assist in and facilitate the supply of information to researchers for enhancing the visibility of their research output and their personal and research teams' profiles as institutions manage their internal promotions and their outward-facing impact globally.

INTRODUCTION: Kara Jones (University of Bath)

Kara Jones (Research publications librarian at the University of Bath and chair for the day) was really the poster-child for a move into this field. From a subject-liaison role she then managed an institutional repository before her current involvement in supporting her institution's academics as authors and Bath's publications database. It was not hard to see the "natural progression" towards working to support the research evaluation process wholesale. Kara highlighted further themes that emerged throughout the day – the "appetite" institutions and academics have for the skills librarians can offer in this area, and from the libraries' perspective the opportunities for advancing their agendas – particularly for Open Access – on the back of their greater involvement in research support.

KEYNOTE: Andria McGrath (King's College London)

Andria McGrath (King's College London) set the scene with a survey of developments over recent years and the ways in which librarians have sought to interact with research evaluation. From the former RAE to the forthcoming REF, institutions have needed librarians' help, knowing of what vital importance the RAE/REF is to their future income. Similarly internal evaluation is ongoing and the data also vital in the awarding of promotions and in appraisal. At King's interest and excitement has been fostered by a series of REF events from 2008-2010 of which a strong element was in addressing the question of how staff were going to skill-up to support research evaluation. Subsequent events at King have looked more closely at which systems were being implemented – if any. In 2009 Symplectic and PURE were heard about, but there was a general innocence almost about what was out there. Institutions' needs, it emerged, were various, with some focussed on wanting to gather all an institution's data on research while others were primarily concerned only with satisfying the data requirements of the RAE. In summary of the progress people were making at that stage, Andria showed how institutions were essentially in need of a single point of data entry while continuing to manage multiple systems for different purposes and/or sectors. An "integrated system" was sought, populated by data from sources that would, inevitably, be separately maintained. Concluding with a final event at King's in June last year, the next big thing on everyone's mind was research impact, and not merely publications' impact, but its practical impact and in new areas such as translational medicine. While collecting data on research impact was seen as ostensibly more

straightforward as it is largely textual in nature, there was at the same time, in June 2010, a growth of awareness across all involved of the importance of research data of all types, bibliometrics *per se*, and the need for systems to collect it.

In mainland Europe the situation is more advanced, so much so that a data standard had been developed for the management of research data – CERIF (Common European Research Information Format). This uses an entity relationship model, comprising objects/entities possessing attributes which can be linked in linking relationships and to which can be added semantic features. (After the workshop, totally ignorant almost of the research data world, I found some of the most recent "tutorial" slide-presentations Andria pointed out as available on the CERIF site really useful in starting to get familiar with CERIF.) For instance, in its semantic layer, the time-stamp fixes researchers' changing relationships as their institutional and research group affiliations change over time. In partnership with the University of Southampton, King's developed a bespoke XML schema from the CERIF standard (CERIF4REF) in the JISC-funded R4R project, intended specifically to allow the extraction of data from an institution's institutional repository (IR) into a CRIS for passing data to the REF, in the format the REF is presumed or likely to require (since the start of 2011 plugins to generate CERIF4REF files from ePrints repository software, Fedora and DSpace have been made available by the project).

So librarians could and should get involved and if not being involved, should try and involve themselves, not least in terms of information resources as many CRISs require subscriptions to services about which librarians will be able to contribute on licensing options, etc. The RLUK also is steering librarians towards greater involvement as it redefines the research library model. Here Andria referred us to the Research Information Network (RIN) report *The value for libraries for research and researchers* and the forthcoming RLUK report *Reskilling for research* which indicates some of the skills libraries will need in the future to support research. Indeed the databases (WoS, Scopus, etc.) from which much of the data is drawn are librarians' bread and butter anyway; it's a case of looking for new ways of interacting with these databases, citation reports that can be generated from them, learning about APIs and how to draw out the data and use it in other systems, discovering what products can help with analysing bibliometrics. Help is available from Thomson Reuters – and the excellent training on offer from MIMAS on Web of Knowledge was mentioned too. Getting further into detail such as the key importance of normalization of the data (by different spans of years; by document types) it quickly became clear training was vital. Andria had seen how academics' interest was propelled by activities like collecting publication data for their ResearchID profile and how significant steps forward were made when research leaders were on side and promoting bibliometrics – e.g. at King's Institute of Psychiatry the Dean had been using baseline tables to show percentiles for citations when head-hunting and evaluating new candidates for research teams. At the start of post-graduates' research librarians could start to make themselves useful giving training on strategies for getting published and where and how to "make an impact" in their early careers. This was a theme that recurred through the day as librarians can relate research impact to the Open Access agenda and the wider context of the shifting paradigms of academic publishing when addressing researchers in sessions that are nominally based on research evaluation. Having demonstrated how the different players were coming naturally together, how the synergies that already existed in their work aided this development, but also how the overall picture of data entry, storage and re-purposing was complex, Andria showed that what academics were ultimately after was a dashboard-style interface to the data that was "super-easy" enabling them to see all the measures they want, constantly updated. This "holy grail" of a research information system was ever more desirable as the data was wanted for a whole gamut of needs from updating personal web pages to completing grant applications. Persuaded research evaluation is librarians' business? Definitely.

Responding to questions Andria confirmed the importance of normalizing data to the world average citation rate, and described the CWTS "crown indicator" – the actual citation rate divided by the world citation rate. It was asked: Should we attempt to force the pace of change with our institutions' research offices? Or let them make their own decisions? Essentially Andria thought we should be in the business of making it easy for them. Keeping staff profile pages up-to-date for instance is notoriously not done; librarians should find themselves in the position of facilitating the use and supply of this data that includes information on grant awards and students taught to promote the best possible picture to the outside world. Without much digging it emerged it was at the University of Leicester that a bibliometrician had been appointed, although not within the Library.

LIBRARIANS' INVOLVEMENT WITH CRIS DEVELOPMENTS: Jeremy Upton (St Andrew's University)

Starting with a timely reminder that HRH Will and Kate Windsor are graduates both of the University of St Andrew's, Jeremy deftly led into a setting of the context for the success story that is CRIS implementation at his institution. In general terms, the developments moved more quickly thanks to the relative smallness of the campus. (And it's true for anyone working where the research office is if not miles away from the library but feels like it this in itself can be a barrier.) St Andrew's physical compactness a given, it was really the library's positive approach to its involvement in the university's research agenda in future combined with what seemed like just a fortuitous "coming together" of related streams of activity – open access, research profiles and impact, digital communication – that made the journey easier and faster. The groundwork for the CRIS implementation came via the development of the e-theses service that in turn grew out of the HaIRST project (Harvesting Institutional Resources in Scotland Testbed) in which St Andrews had been a consortium member. Vital to the early involvement here was getting out and talking to people outside the Library – staff in the registry office, supervisors, graduate schools. Human interaction – it seems so simple, but at St Andrews there is no doubt of the part it played in bringing results as Jeremy charted those early interactions up to the present through the key post of repository assistant to the staff of the PURE team.

St Andrew's developed a Research Expertise Database (RED) and has worked hard to bring that together, both through personnel and technically, with its IR, prior to replacement of the RED with the PURE CRIS. People working in each project clearly had much to learn from each other's experience. Relationships grew stronger and academics increasingly saw library staff as their natural partners in the process. From here institutional restructuring and the procurement of the PURE system could start from a very solid foundation. Jeremy pointed up a chief catalyst for change was the 2008 RAE for which librarians and repository staff members were essentially obliged to work very closely together to provide the requisite data. Librarians of course understood how the data was structured and recognition was made that bibliographic knowledge management was the library's area of expertise. Rebranding into a "PURE team" there is now a strong web presence for research at St Andrews in the blog. A portal supplies a public-facing platform for the publications stored in the IR.

St Andrews, Jeremy reported, had more work to do encouraging academics to deposit in the IR via PURE, but the benefits on all fronts continued to deepen commitment from all involved in the project. Library staff were actively involved now in promoting the OA agenda on the back of research depository training and in turn they were learning much more about the research St Andrews was currently engaged in. Graduates and the public could see how the university's subject-based websites were populated by data (on theses and research papers) from the IR and the repository team could disseminate information on the university's internal publishing of its own research via the OJS (Open Journal Systems). These and other advantages following the implementation of PURE are summarized in *Increasing uptake at St Andrews – strategies for developing the research repository*. Standing out as a special achievement is the doubling of full text deposits in the IR, growing by 52% between June and December 2010. With the REF on the horizon the joined-up but apparently organic approach at St Andrews left Jeremy almost wanting a more negative paper to follow his to counter-balance it, but the only issue that posed some challenges was the interface and communication with the (Scottish national) consortium responsible for the IR. Questions led into the use of the CRIS as a full-text delivery system itself and the potential for larger-scale restructuring within institutions as librarians roles and skills changed to accommodate research data support and evaluation.

SYSTEM DEMONSTRATIONS – PURE; InCites. SURVEY RESULTS

Following Jeremy Upton's retrospective take on having implemented the PURE CRIS at St Andrews and the contexts that led up to the present, it was especially useful to see PURE from a very different stance, close-up and personal, as implemented at the University of York. Here the library had not been involved in the procurement of the CRIS until relatively recently and now taking on a support role for PURE had been more sudden than anything organically introduced over an extended period of time. Sue Cumberpatch (Academic Liaison Librarian) described how the initial bulk upload of data had been performed by Atira (the company that designed and owns PURE), followed by individual academics adding their own data or paying students to add it, which necessarily raised issues and

concerns over quality control and the prospect of data clean-up in the nearer future. Now York was aiming to go live from September/October 2011, the first university outside Scotland to take PURE (followed by Leeds).

Sue gave a system demonstration of PURE, indicating the different options for adding research data – (i) manually via a template (the method used by the student workers at York); (ii) by input from an online source (PubMed; arXiv.org; Web of Science); (iii) by import from bibliographic databases (e.g. MEDLINE). With manual input there were serious concerns over validation and curation of the data (and this was viewed as an option not to be preferred in the long-term). Input via online sources actually worked well, but chiefly only for the sciences; as soon as social sciences and humanities research was sought, this method fell down. Using BibTeX files created from EndNote for imports from databases had not been straightforward and the BibTeX files had required some tweaking. Convergence of IT and Library services at York had meant, though, that the technical help was on hand to get this to work. All in all, researchers did need librarians help in a big way – bringing in data; setting up scans for and adding new outputs of data; dealing with discrepancies between links to e-print versions of papers in the IR and links to published versions in WoS etc. York has started help pages for using PURE that are publicly available. Going forward, there were potential issues ahead with the integration of different CRIS systems chosen by institutions sharing a common IR – the White Rose Research Online (WRRO) consortium – as Leeds and York have selected PURE but Sheffield uses Symplectic.

Andria McGrath then summarized how InCites from Thomson Reuters was used at King's College London by demonstrating the creation of a data set of researchers working in the field of addiction, an author set extracted from the King's CRIS. Andria explained the importance of the 1% percentile and showed graphical illustrations of the proportions of authors in an author set above and below the global average. It quickly became clear that InCites could reveal an immense amount about an institution's research standing but the data had to be a very good set that is sent to Thomson Reuters; it had to be reliable and so populating the CRIS with researchers' publications is vital. Andria could also show in InCites how your data could be used to make comparisons with rival institutions and against different subject classification schemes. Once tables are converted to graphs the data obviously becomes much easier to read and here InCites can uncover what that data really says about an institution's impact relative to a subject area over a given period of time. This information is used subsequently in preparing research reports and for supporting bids for grant funding.

A survey was run before the workshop inviting institutions to respond to a series of questions on the implementation and use of CRISs. Disturbingly still perhaps half of 50+ respondents had no CRIS at their institution. And the vast majority do not subscribe to InCites. There was a trend also that emerged from the survey that showed that decisions on acquiring a CRIS were taken at a very senior level possibly more remotely from practitioners who were trying to build partnerships across institutions. However there was some consensus products like InCites and bibliometrics generally were becoming more important.

THE ROLE OF THE LIBRARIAN IN THE RESEARCH EVALUATION PROCESS: Kate Bradbury (University of Cardiff)

Kate Bradbury began with an effective tag cloud displaying the concerns of academics around research data. Of course in centre place is the REF 2014, but Kate stressed it was not the REF alone that was the driver for institutions to get immersed in supporting research data and evaluation. In fact the drivers are many, from the THES world rankings - for which citations contribute a large part of the ultimate results – to the dependency of grant applications and awards and internal promotions and departmental performance.

The librarian supporting research evaluation is now fielding questions from researchers – Why bother with looking at citations if the REF isn't going to use them? - Is it going to be worth our while to subscribe to InCites or SciVal? - What can these services do that our current databases can't? Kate highlighted areas of knowledge and expertise we could all develop as working in the research evaluation process revealed where researchers' needs were directed. Collaboration and partnership in research teams had been commonplace for years, but it is ever more relevant to expose which institutions you are collaborating with, and which countries in the world – both currently and in the

future. From the institutional level to the individual researcher wanting article level data, the calls for the librarian to unpack the analyses it is possible to perform in Web of Science and Scopus are growing apace. Perhaps unexpectedly there were opportunities here for working with suppliers on developing their products and Cardiff had worked successfully with Elsevier and Thomson Reuters on, for example, reconciling research affiliation data which can be problematic as institutions undergo changes of name, never-mind researchers changing their individual institutional and research team affiliations.

Librarians were, as was later noted in the workshop, a neutral space to which the skills of bibliographic data management key to the research evaluation process could be trusted. The correctness of data was vital for its re-use and here and in search skills librarians had a central role to play. As well as preparing and producing the requisite citation reports and analyses required at Cardiff, Kate delivers bibliometrics training for groups and individuals. She described how the IR manager's role had moved from the back office to liaison with multiple stakeholders in the university. Again there were opportunities for advocacy of the Open Access agenda via this training and liaison. Pointers were given for the future. Not least is for the librarian to gain an awareness of the research landscape in terms of the forthcoming REF (e.g. the Research Information Network made surveys of the needs of researchers that can be followed). Get a better understanding of what InCites could do to enhance research evaluation in your institution and the limitations you are placing on your data without it. And get help from the experts – learn those shortcuts in Excel and explore what it can do for you.

SUPPORTING BIBLIOMETRICS: Jenny Delasalle (University of Warwick)

In her overview of her role "supporting bibliometrics" as Academic Support Manager (Research) at the University of Warwick, Jenny Delasalle recommended a "sceptical and informed" stance as so important to supporting research evaluation. It was essentially for researchers to decide how to deploy the information the library could provide on strategies for publication and measuring its impact, and not for the library to determine researchers' choices on what to do with the information. In a sense this stance could be at odds with researchers' hunger simply to be told what they thought they needed to know, but the critical approach to delivery of information in a complex area could, over the longer term, be highly influential in steering policy and practice. In guiding researchers at Warwick Jenny described how training should comprise a critical appraisal of different databases and be wide-ranging in educating researchers in different sources of data for ranking.

In terms of preparing data for the REF 2014, researchers needed to be discriminating in how a set of papers were defined for measurement; how papers can be scored against different totals and in particular how a researcher's h-index score will vary dependent on the data sources used to calculate it. There were questions raised in discussion with academics over scoring that could obscure, but should reveal, whether it is better to reach that h-index ranking from a single paper cited a hundred times or from ten papers cited ten times and in the sciences researchers should be aware of the different weighting given to titles in different data sources – the Academic Ranking of World Universities (ARWU, aka "The Shanghai Rankings") for instance will give more weighting to papers published in *Nature* and *Science*. For researchers on the cusp of disciplines different strategies for measuring impact were called for (e.g. in management studies); clearly comparisons within single disciplines were more meaningful. In looking forward to the REF 2014 Kate noted the potential tension between reliance on data from journals with the highest Impact Factor and the need to show socio-economic impact and research with the widest possible dissemination.

GOOGLE SCHOLAR: CAN IT REALLY BE USED FOR BIBLIOMETRICS? – Isobel Stark and Michael Whitton (University of Southampton)

Google Scholar – as everyone knows – is the first port of call for most academics, so with the REF 2014 on the horizon and questions being asked on using GS for it, we've got to answer them. Even with the apparent retirement of investment by Google in Google Scholar and the increasing edge Web of Science will have with the advent of the inclusion of scholarly monographs, Isobel made a strong case for GS as a source of citation data and for research evaluation. Where WoS was limited – in

certain fields of research and in the cross-over between different subject areas, Google compensated significantly – e.g. in archaeological computation (a strong research area at Southampton) GS was excellent but WoS not good for either archaeology or computation. As for medics there is evidently no competition; astonishingly GS got 5,000 more uses than WoS over the last year by medics at Southampton calculating their h-index. The quick and dirty approach is fine for them, is all they want. And indeed grant applications do not ask for an h-index built from any particular data set. What's more Google has done much to improve the quality of its data in GS which has been such an issue since its launch in 2004. And it is completely free.

Isobel then probed deeper into GS and the discoveries were fascinating. While the data is less ropey now, it has to be said that it remains utterly opaque how the data is gathered. There is no transparency over which publishers Google is in discussion with for their data and no clarity about the gaps in coverage which are crucial for any measurements on impact over periods of time. With no human intervention with the data, it is far more time-consuming to disambiguate names of researchers to get an accurate data set (the Sofaer brother, sister and father – all academic researchers – was a nice case in point). With the h-index the picture is intriguing. Although overall the metric is usually higher with data derived from GS than it is with say Scopus, it emerges that there is a wide variation within GS that appears to be linked to disciplines. As the h-index comes out higher, the same, or lower consistently in comparison with WoS and Scopus it averages out, but the variables have riders with them. Computer scientists' scores, for instance, come out lower in WoS probably due to the fact that their research is published in conference proceedings that are commonly not re-issued as journal articles so miss the h-index calculation entirely. In the social sciences h-index scores in GS are derived also from publications in scholarly monographs that (at least now) WoS will omit. The indexing of articles in PubMed by GS similarly has an impact which looking for instance at neuro-surgeons' scores boosts the results significantly. Looking at the scores of individual researchers at Southampton, Isobel could substantiate these effects by comparing h-index scores in WoS against those from GS for the same academics. Professor Jane Falkingham, working in the field of osteoarchaeology was on the cusp of multiple disciplines, publishing in both cultural studies and medical journals (GS 23; WoS 7). Other researchers in medical sciences were similarly scoring higher in GS thanks to the inclusion of scholarly monograph publishing in GS and PubMed article data. Professor Nigel Shadbolt, a computer scientist, similarly scored higher in GS; he is publishing in conference proceedings with occasional forays into IEEE journals. Dr Simon Coles (Chemistry), on the other hand, scored 22 in GS and 36 in WoS. Isobel attributed the cause of this to the pulling out of GS of ACS publications; following ACS' lawsuit against GS (for its trademark violation of and unfair competition with SciFinder Scholar) the only data indexed by GS of ACS now is likely to be scrapings from openly available material on the ACS website.

All of which is almost as much an education in WoS as it is in GS, and led into discussion of whether other data sources could or should be used for citation measurements (e.g. PsycINFO). See Southampton's help pages.

There was discussion around topics between the workshop delegates and the speakers:-

- Sharing information. Two relevant maillists were noted: LIS-RESEARCHSUPPORT and LIS-BIBLIOMETRICS and the Association of Research Managers (ARMA)
- There'd been discussion and demonstration of deriving data from say WoS or Scopus or Google Scholar. Should we try to combine data from all 3? Would this be best practice? First off, this would involve a colossal amount of extra work, however desirable. Second the techniques used to arrive at final result sets would have to be replicable both in subsequent evaluations at a single institution and against other institutions for benchmarking.
- APIs developed for extracting data from WoS and Scopus could over time allow such comparisons / combinations of data to be made easier
- Thomson Reuters training was strongly recommended (training includes techniques for normalization (by year; by document type))
- WoS supplies an "expected Impact Factor" for journals less than 2 years old in publication before which time titles lack an IF.
- What work did the repository assistant at St Andrew's chiefly do? Improving quality of bibliographic data coming in from data collected from academics. In general there was a massive under-estimation of the amount of time and resource that needed to be put into this type of work for fulfilling research assessment satisfactorily

- A review of what libraries brought to the process: bibliographic knowledge; advocacy; training; standards; tools; quality control – all of which amounts to a trusted brand and a combination of offerings to be found concentrated nowhere else in one team in most institutions

- Quite laborious and difficult to manage people's desire just to be told what they need to do – e.g. not to be informed of the details and vagaries of calculating the h-index. Balancing that with retaining the library's "neutral space" – neither purely administrative nor purely academic – providing an informed overview of the pros and cons.

- There was discussion over the preferences of academics in different disciplines for subject specific data repositories over generalized portals offering access to agglomerations of data. In turn different repositories evidenced better and worse practices for the storage of metadata on the research content in terms of its re-use and long-term preservation (see: eCrystals: Crystallographic Repository at Southampton as a model)

- For storing full text it was argued an IR is better than a CRIS, but ideally should be used with a CRIS as the deposit mechanism (institutions showed a wide variety in their configurations of IR and CRIS implementations). In continental Europe there was more of a trend for use of the CRIS as an institution's IR also. In the UK there may be more of a separation of powers whereby the IR is run out of the library and the CRIS resides with research support (located outside a library). As we'd seen – in some institutions the CRIS implementation could bring these teams together to form a single, dedicated service. As libraries routinely interface with all levels of research they were well placed to take opportunities to develop groups for distinct user needs or bring those constituencies together.

- There was concern over the display of data via the CRIS and manipulation of IF scores. Due to the dominant focus of the journal IF some "guerrilla action" had been seen where citations had been deliberately made to titles in order to boost their IF score so as to influence personal rankings.

- In thinking about forging relationships a growing trend for publishers to bypass the library entirely and go direct to academics was noted.

- MyRI – Measuring your Research Impact is a new project starting up in Ireland :-"an open access toolkit to support bibliometrics training and awareness".