
Conference review

'Subject repositories: European collaboration in the international context' – a conference organised by the Nereus Consortium

British Library Conference Centre,
London
28–29 January 2010

Conference reviewed by Antony Brewerton,
*Head of Academic Services,
University of Warwick*

In issue 42 of *SCONUL Focus* Hans Geleijnse and Pete Williams heralded the launch of the European Union-funded project NEEO: the Network of European Economists Online. Explaining the project succinctly, they stated:

*'NEEO's aim is to bring together Europe's top economics research and make it available through a new multilingual portal called Economists Online.'*¹

That was in September 2007.

Whizz forward to January 2010: the project is almost complete and Economists Online is ready to be launched at the conference 'Subject repositories: European collaboration in the international context', organised by Nereus, the body behind NEEO.

Jean Sykes, Chair of Nereus, welcomed us all to the conference, which, as its title suggested, was

a truly international affair. The conference had attracted 174 delegates from 23 countries across the globe. These 174 included librarians, IT staff, project and repository managers, data specialists, research officers, academics and researchers, publishers and funders.

Part of the reason for the conference was to explain the purpose and gestation of Economists Online and to see the repository formally launched. The wider aim of the conference was to put all of this into the broader context of the whole institutional repository, subject repository and open access movement.

ECONOMISTS ONLINE

Hans Geleijnse, the NEEO project director and president of LIBER (Ligue des Bibliothèques Européennes de Recherche, the Association of European Research Libraries), began by looking at Economists Online in the context of European subject repositories.² The objective of EO has been 'to improve the usability, global visibility and management of European economics research'. The project began with 16 member institutions, all coming from different starting points. Some had well-established institutional repositories from which EO's data would be harvested. Others were new to the repository world. A wide range of repository software was being used: DSpace, EPrints, ARNO, Fedora, as well as home-grown systems. In fact, the only unifying element – apart from a desire to make the project work – was that nobody had much experience of dataset management.

Despite the challenges that all this brought, EO more than exceeded its aims. It has attracted publication lists from 836 academics (having sought 500); 73,506 metadata items (instead of the target of 50,000); 20,500 full-text (historic content) records (instead of 15,000); 7,982 current content full-text records (not 7,500) and six new partner institutions (two more than the original target). The only area where the project has fallen short has been with datasets, with just over 100 secured instead of 160.

But why use EO? In many ways it has entered a crowded market place, which includes Econstar, SSRN (Social Science Research Network), RePEc (Research Papers in Economics) and AgEcon Search (Research in Agricultural and Applied Economics). Hans finished by considering the unique selling points of EO:

- It is based on institutional repositories.
- It is supported by the institutions behind these IRs and the academic community.
- It promotes open access.
- It includes the publication lists of over 800 leading academics.
- Its content includes post-prints, books, book chapters and datasets.
- It has high-quality metadata.
- It has a multilingual search facility.

EO also aims to collaborate, rather than compete, with other providers and is already building relationships with many of the key players in the field.

Nicholas Barr, professor of public economics at LSE (London School of Economics), was then called upon to launch the database. In 1948, he reflected, an economist could claim to know 'all of economics': Keynes would read all his journal subscriptions from cover to cover and absorb all the knowledge available. Today the amount of information produced is unmanageable. Researchers need something to make it manageable and EO is one such tool, providing a one-stop shop for quality information and allied support, Nick declared.

We were then treated to an audio-visual countdown giving key figures about the database before EO was declared live:

- 940,955 bibliographic references
- 28,482 full-text documents
- 836 publication lists
- 100 datasets
- 23 partners
- 4 languages
- 3 continents
- 2.5 years in development
- 1 portal: Economists Online.³



Figure 1. Nicholas Barr launches Economists Online

The next day *François Cavalier*, library director, Sciences Po (Fondation nationale des sciences politiques), outlined his experience of the NEEO project, as a project member. Sciences Po was new to repositories so this proved a real learning curve and François reviewed some of his steps on the way. So was it worth it? It was hard work but Nereus provided the opportunity to do new things with the assistance of a wide and supportive network, and at the same time to push the open access agenda further forward.

Following François was *Wietske Sijtsma*, NEEO project manager, Tilburg University, presenting on the key features of Economists Online (with a demonstration of its functionality) and the lessons learnt from the project. To this end, Wietske's paper included inputs from the previous project manager, *Vanessa Proudman*.

The main lessons learnt from the project were:

- Get the researchers involved ... if you can: regular surveys and focus groups were carried out but it was difficult to get inputs from busy people.
- Researchers do not complete questionnaires.
- Researchers are not interested in beta projects – they are only interested in the finished product.
- If you do get ideas from your researchers – use them.
- Getting content is a big issue – it won't just come to you: you need to go out there and sell the repository (which is a new approach – requiring new skill sets – for some library staff).
- Getting new types of content is difficult – datasets were a particular challenge.
- Measuring success is important – quantitative and qualitative targets need to be set.
- Once your repository is established, make sure it is sustainable – consider staffing, structure and preservation issues.

Some of these issues have been successfully addressed. Others are being worked on. All of them provided considerable learning opportunities for both the project managers and the NEEO partners.

SUBJECT REPOSITORIES

As well as focusing on Economists Online, its launch and an appraisal of the project so far, the conference also looked at the wider theme of insti-

tutional and (especially) subject repositories in the context of open access developments worldwide.

Javier Hernandez-Ros, head of the access information unit, directorate general for information society and media, for the European Commission, gave the view of scientific information from Brussels and provided an overview of recent developments in the open access movement. *Dale Heenan* from ESRC (UK Economic and Social Research Council) and *Neil Jacobs* from JISC (Joint Information Systems Committee) explored the linkages between institutional repositories (IRs) and subject repositories (SRs), ways in which they can work more closely together and the roles to be played by funders, researchers and their host institutions.

Cathrine Harboe-Ree, University Librarian at Monash University, then took us to the other side of the globe to consider the centralised approach taken in Australia. Unfortunately, our speaker had lost her voice so Jean Sykes came forward to speak to her slides, which covered: (a) key funding and policy initiatives; (b) the work of Monash University in this context; and (c) the issues and challenges Australian colleagues face.

One area where our colleagues are less challenged than we are seems to be in funding: the Australian government has invested \$Australian 400 million over four years for the sector to develop its e-research infrastructure.

Patricia Renfro, deputy University Librarian from Columbia, then flew us over to America for 'Subject, institutional and other repositories: a US perspective'. She reviewed the five major subject repositories (arXiv (for physics, mathematics, computer science, quantitative biology, quantitative finance and statistics), CiteSeerx (which focuses primarily on literature in computer and information science), RePEc, SSRN and PubMed Central (the US National Institutes of Health's free digital archive of biomedical and life sciences journal literature)) and their shared characteristics. All are tightly linked to their subject discipline, which is important because researchers identify themselves with their discipline before their institution. Researchers see the benefits of SRs more readily than they see benefits in IRs. To help address this, some universities (for example, Rochester) have produced a suite of services assisting authors to publish or self-publish their work, to provide added value to their IRs and to attract researcher interest. Another route involves

linking IRs to the more attractive SRs, just as EO has done.

Clifford Lynch, director of the Coalition for Networked Information, rounded off day one with a session in which he said he would 'speculate ... irresponsibly'. Sustainability of SRs was raised. IRs – and their host institutions – are not likely to disappear but SRs (and their funding) are less secure. Data management came up again as an issue. Again, IRs may be starting to collect data but curation and exploitation may be more appropriate in the SR environment. Another theme was selection of content. IRs' quality control is based on the institution's staff selection policy. For SRs this is less straightforward. Peer review or an editor may be a way forward but this is hardly scalable. Again, perhaps the answer lies in SRs riding on the back of IRs. All this pointed towards more projects like NEEO across various disciplines.

Christian Zimmermann from the University of Connecticut kicked off day two with the view from RePEc. This was a lively session that really got the day off on the right note. I especially enjoyed Christian's reflections on relationship-building with various stakeholder groups. He finished with a list of key messages applicable to all repository builders:

- You need a core group of believers to get things started.
- Ensure you get no institutional hindrance – in some areas powerful subject bodies/associations can prove difficult.
- You need people with some 'computer savviness'.
- Just do it – stop talking about it, just do it!

Day two also saw the opportunity for delegates to attend one of six parallel workshops (based on EO work packages) covering key – and generic – repository themes:

- content-recruitment issues
- infrastructure and interoperability
- intellectual property rights
- multilingual tools
- datasets
- usage statistics.

I attended the session on datasets, led by *Thomas Place* from Tilburg University. Having found securing data a particular problem, I was keen to hear about the approaches taken by other colleagues. It was good to pick up tips but also reassuring to hear that we all face the same challenges.

SUMMING UP: KEY THEMES FOR THE FUTURE

So what did I take away from these two days?

Data-curation was one of the biggest themes – if not, indeed, the biggest – of the conference for me. As François Cavalier put it, ‘Data management is the new frontier for libraries.’ It is also problematic, as Hans Geleijns pointed out. The reasons for EO not reaching its targets here were:

- Economists’ datasets tend to be commercially produced and not created by the researcher.
- There is a desire to see other researchers’ data ... but a reluctance to share it.
- Intellectual property rights are more complex when it comes to data compared with published outputs.

These comments were amplified in the datasets workshop.

Another challenge was thrown down for us by Nick Barr, who said he wanted ‘smart’ technology in the repository world. ‘Smart’ developments mean the cleverness goes into the product so that the end user does not need to be clever to use it. So a smart repository would mean that Nick could ask for articles that would interest him ... and the database would respond appropriately. Start building your end-user profiles now ...

Also important for me was the potential relationship between IRs and SRs. IRs seem to be the way to get the content in but SRs seem to be the way to get the content out (and thus attract content in the first place). Perhaps we need to build the repositories in our institutions and then work with our discipline colleagues across the globe to build the add-on SRs to really exploit the content and create a virtuous circle of development.

A challenge – or perhaps an opportunity – is Mendeley (www.mendeley.com), as raised by Patricia Renfro. This ‘iTunes for research papers’ is a social networking repository employing drag-and-drop technology that has over 100,000 users, 8 million articles and is doubling in size every twelve weeks. Perhaps the future isn’t IRs or SRs but this?

Whatever route we take – IR, SR, data repositories, Web 2.0 solutions or something nobody is even thinking about at the moment – another comment that stayed with me came from Hans Geleijnse, who finished his introduction by saying, ‘Libraries can be catalysts for change.’

These – and other themes – were neatly bundled up by *Paul Ayris*, vice-president of LIBER and director of UCL (University College London) library services, in the concluding session. Paul’s top ten countdown of messages were:

- 10 The conference had given a good summary of the differences between IRs and SRs.
- 9 EO needs to consider widening participation – not all excellent economics research is published by researchers at the most highly rated higher education institutions in Europe.
- 8 Back-digitization of core papers would be a useful next stage.
- 7 Mandates are important and Nereus could provide advocacy work to show their benefits.
- 6 Funding is always an issue – and we probably won’t be as lucky as our Australian cousins – but Nereus could do research into funding opportunities for repositories.
- 5 The issue of sustainability was raised various times during the conference, with IRs in a stronger position than SRs.
- 4 EO provides digital preservation via an agreement with the Royal Library in The Hague – this is an added value that other SRs could pursue.
- 3 EO has encouraged library staff to develop new skills in advocacy and marketing.
- 2 Repositories are not just about content but about services, like building virtual subject communities and other value-added offerings around the repository, such as developing new e-publishing models.
- 1 Unsurprisingly, data management was in at number 1 and was Paul’s final point: this is a huge issue – and an even bigger opportunity – for information professionals.

This was an excellent and thought-provoking conference. To me it didn’t feel like an end of a project but the start of something far bigger.

NOTES

- 1 Hans Geleijnse and Peter Williams, ‘NEEO and Economists Online’, *SCONUL Focus*, 42, Winter 2007, pp33–5
- 2 The slides for the papers presented can be found on the Nereus web site at <http://www.neeoproject.eu/conference-speakers.html>.
- 3 The portal is available at <http://www.economistsonline.org/home>.

All web sites accessed February 2010.