**Counting on *Open*DOAR**

*Peter Millington, Centre for Research Communications, University of Nottingham, UK*

*Open*DOAR has been transformed during the past year into the repository registry with the most accurate record counts. How has it done this? Bear in mind that it has to monitor over 2250 repositories, and a third have no machine interfaces such as OAI-PMH.

The entries in *Open*DOAR have always been quality controlled and deduplicated. There was, however, a problem with record counts, which were updated infrequently, if at all. This Achilles Heel was tackled in early 2012, and now 95% of entries have accurate record counts, regularly updated. There are over 142 million records in *Open*DOAR’s listed repositories at the time of writing.

The main established methods of capturing record counts in the past have been:

(a) To inspect the repository manually, or

(b) To use OAI-PMH to harvest and count records or record identifiers.

The manual method was prohibitively labour intensive, while OAI-PMH proved to be inefficient, unreliable and sometimes downright inaccurate.

While not rejecting these methods completely, *Open*DOAR has developed a suite of harvesting methods, many of which might be described as “screen scraping” – for instance, extracting the “*n*” from DSpace’s “Now showing items 1-20 of *n*” displays. This approach has obvious drawbacks, such as the need to cope with website redesign, but it has nonetheless generally been found to be much faster, more accurate and more reliable than OAI-PMH.

The twelve harvesting methods we now use reflect the different ways in which repositories show data and indexes on their sites. I will briefly outline these methods and their respective pros and cons, and will use the repositories and platforms that handle record counts best as case studies – naming and proclaiming.

We have also been able to count the number of full texts in some repositories, although like everyone else, we find this an on-going difficulty. I will outline some possible solutions to this problem.

Does *Open*DOAR still not have a record count for your repository? Is the record count still out of date? I will give you tips on the simple things you or your software providers could do to provide the information we require, and improve the profile of your repository into the bargain.