

Institutional Repository Checklist for Serving Institutional Management

Leslie Carr, Wendy White, Susan Miles, Bill Mortimer

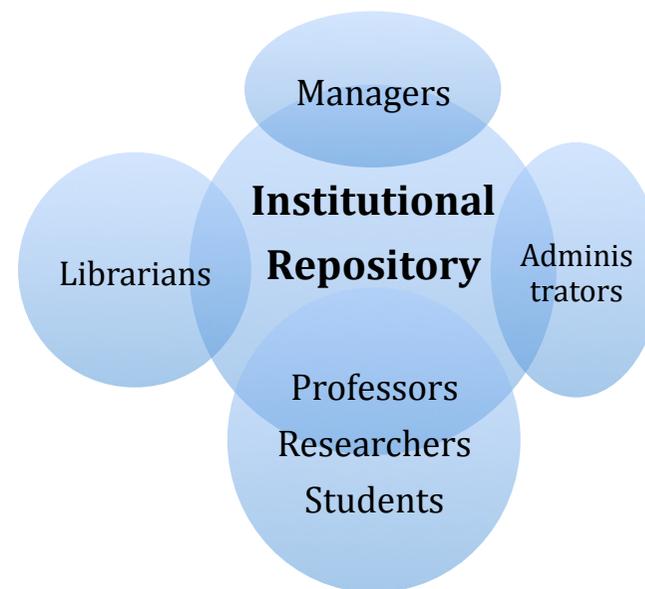
This document is a For Discussion draft that came from the Research Assessment Experience session of the EPrints User Group at Open Repositories 2008. Comment is invited from managers of all repository platforms to share experience of the demands of the processes involved in supporting research assessment at an institutional level.

The aim of institutional repositories has focused on serving the interests of faculty – researchers and teachers – by collecting their intellectual outputs for long-term access, preservation and management. However, experience shows that in order to attain the engagement of the faculty, it is necessary to obtain the support of the institutional management. Even more fundamentally than that, the support of the senior management ensures the long-term financial and political survival of the repository and of its objectives for the near- and medium-term future.

Most universities have a three-fold mission statement that encompasses advancing knowledge through research, communicating knowledge through education and helping to apply knowledge within society. An institutional repository can help to deliver on each part of the mission, but that aspirational level of engagement isn't enough to win institutional-level support. A repository has to demonstrate a measureable and effective contribution to current management agendas and concerns – *e.g.* research management, raising international profile, improving citation metrics. Such contributions are achievable, but only if the repository fulfils a number of criteria that are in addition to its usual library-faculty roles.

The success criteria in the following table is created from discussions with managers of repositories with established roles in reporting to institutional management. The numeric risk factors in the second column reflect the potential consequences of failure to deliver according to an informal nautical metaphor: (a) dead in the water, (b) quickly shipwrecked and (c) may eventually become becalmed.

	Criteria	RF	Capability	Risk	Comments
1	The repository must uniquely identify each local (institutional) author for each item.	a	Accurate lists of publications can be produced for each individual.	If the authorship of each item is determined only by human-entered names then metadata inconsistencies will inevitably make it impossible to produce an authoritative list of all the items authored by (or edited, conducted, produced or contributed to by) an individual.	The id could be an email address or staff id, if not reallocated after the staff member leaves. Ensuring persistence may be a policy issue for Information Services or the Personnel or Human Resources department.



2	The repository must contain a complete, accurate and updateable list of all the academic staff, faculty, researchers and professors in the institution.	b	Complete publication lists can be produced for all staff.	Publication reports for individuals may be missed and accurate statistics of repository coverage are impossible.	This information may come via the Active Directory or LDAP interface that the repository uses for authentication.
3	The repository must track the affiliation between individuals and their departments or research groups.	b	Accurate publication lists can be produced for departments or research groups. Publications can be accurately credited to the appropriate institutional unit and reallocated when management structures change.	When the whole University restructures, or departments/groups split/amalgamate, or individuals extend or alter their affiliations and responsibilities, it is impossible to automatically re-allocate the items to new 'communities'.	Item affiliation is a policy-determined multivalent combination of static affiliation (declared upon deposit) and dynamic affiliation (determined by the current affiliation of the author or any other contributor). See also #10
4	The repository must be able to handle content-free items	a	Lists of faculty works need to include some bibliographic records, e.g. for decades-old work or for rights-blocked material.	If full texts have to be deposited in the repository, credit may not be obtained for items where the full text is unavailable.	
5	The repository must be able to manage dark items	c	Internal management processes can operate on private or sensitive data that is not shared with the outside world.	Some research falls out of the remit of the repository. Parallel systems have to be implemented.	Although important, relatively few items fall into this category.
6	Reports generated by non-system staff	b	Detailed reports and breakdowns of research or teaching outputs must be possible from the web interface	Relying on command line functionality makes the system inaccessible to institutional administrators	Report details and formats change frequently and must be flexibly accommodated.
7	The repository must provide quality assurance processes that can handle high throughput	b	Adding value to self-archived items, e.g. to conform to externally imposed requirements		
8	The repository must be able to support high-throughput deposit processes	b	The depositors' tendency to delay until a deadline means that short-term throughput peaks of several hundred items per day must be accommodated. In the steady state, up to 50 items per day	Large backlogs form. Deadlines are missed for administrative data returns.	Any problem from initial adoption will disappear over time as the repository moves to a steady state, although seasonal fluctuations will continue.

			might need to be handled (1000 research active staff, 1 item per month, 20 working days per month).		
9	The repository should have a policy for dealing with 'very low quality' items deposited by schools or individuals.	b	Items with very bad metadata can clog up workflows; the repository needs to limit its liability in such circumstances to ensure an overall level of service, especially when delivering to deadlines.	Large batch imports from databases or services with low quality control or low metadata requirements demand an inappropriate level of input from, library staff.	Experience indicates an 80/20 split between acceptable and awful deposits. The majority of the 20% come from bulk/batch imports from legacy systems.
10	The repository must support metadata fields to tie its records in with other information systems (e.g. funded projects databases or citation reporting).	c	Other administrative systems can combine repository data for regular management reports.	The ability to 'mash up' or perform cross-domain reporting and analysis of administrative information (demonstrating effectiveness or impact for research funders) may be compromised.	Not all of the institutional research information will originate with the repository, or be ultimately managed there.
11	The repository should involve (and should be involved by) senior management and administrative committees to guarantee institutional embedding.	a	Research leadership and strategy can evolve and be informed by the repository.	Faculty support declines because the repository is sidelined by new management initiatives and other systems that replicate its core functionality without providing document and data curation.	
12	The repository needs a strategy for maximising "full text deposit" along with metadata records.	b	A balance between 'institutional service' and 'individual service' can be established. Library goals of Open Access and Preservation can be maintained and not swamped by the demands of management / administration.	Record keeping requirements may deprecate deposit of full texts; a repository can quickly become a bibliographic database.	Management functions may emphasise data analysis rather than OA or preservation or any form of document/data gathering. This was true for the UK's RAE process, but may change for the REF.
13	The repository needs to support the practices and assimilate the legacy systems of fiercely independent departments and schools.	c	New functionality and better service needs to be offered to Departments who have Their Own Ways Of Doing Things.	Departments may feel a good deal of loyalty and ownership towards their own home-made databases and web portals. This may result in antagonism to the repository, especially if it does not seem to offer any advantages.	Some institutional constitutions make mandates or coercion unlikely in the foreseeable future.

