



Functional Checklist for Digital Repositories in the Research Quality Framework (RQF)

Adrian Burton

This publication is made available by the Australian Partnership for Sustainable Repositories as one of its “Helpful Documents” series issued in association with its work on RQF and Repositories.

www.apsr.edu.au/rqf/http://www.apsr.edu.au/rqf/checklist.pdf

July 6, 2006

Functional Checklist for Digital Repositories in the RQF

Table of Contents

1. Background and Qualification	2
2. An RQF Checklist	2
2.1. RQF Data Model	3
2.2. Complex or Non-text Items	3
2.3. Work Flows	3
2.4. Group Management and Access Management	4
2.5. Communication	4
2.6. Liaison with the Research Office	4

Draft v 1.1 dated 22/03/2006

1. Background and Qualification

From the institutional perspective the planned introduction of a Research Quality Framework will have wide ranging ramifications for research, research administration, and research information infrastructure. This document concentrates on the likely issues for the third area, research information infrastructure. Other issues, such as the role of the research office, the formation of research groupings, institutional research strategies, etc, are out of scope for this paper.

All we can currently say about the RQF must be carefully qualified because the details have not yet been finalised. This may work in our favour; it may still be possible to effect the nature of the final RQF implementation. At this point it would be sanguine to acknowledge that the we don't know how centralised the RQF will be. Will the RQF be running its own mega repository of "submitted" items? Or will "submitted" mean submitted to the institution's repository and the reference submitted to RQF? The latter one would think, but then will part of the RQF process involve DEST harvesting directly from the repository or will the university's research office mediate?

We take as a given the importance of an institutional repository in the process of information management of the RQF process. The introduction of the RQF is a strategic opportunity for institutional repository managers and developers. It is a chance to make the IR even more relevant to the critical mission of the university. That said, it is still possible that the initial RQF will not involve digital repositories at all and will run on hard copy submissions.

Assuming that most universities choose to manage their assessable research information through institutional repositories, the repository software and the accompanying repository services will be called upon to support the institution's response to the RQF. The required functionality for such support is discussed herein.

This paper only tries to propose an indicative checklist of the functional requirements of a repository wanting to support RQF. Further discussion will elaborate upon this list, which is offered at this stage to stimulate discussion.

2. An RQF Checklist

The checklist for institutional repositories in the RQF might look something like this:

1. develop and support the RQF data model
2. support complex or non-text items
3. facilitate work flow for

- depositing academics
 - university administrators
4. manage groups and access
 5. enable communication and automated reporting
 6. liaise with the Research Office

An institutional repository service wanting to ensure that it has the right tools to support the RQF will want to tick off functionality in these six areas. These six areas are discussed in more detail below.

2.1. RQF Data Model

It is obvious that the research items to be collected through the RQF must have appropriate metadata. The metadata must be defined by a schema to ensure that the systems are collecting the required fields. Appropriate vocabularies to describe all the elements needs to be developed to ensure interoperability with DEST and between institutions. The process will need to intersect carefully with existing models such as the RFCD codes. Collections of items (not just the individual items) will also need to have appropriate descriptive and technical metadata.

A common (cross-sectoral) approach to these issues will not only be efficient in the short term, but will also be increasingly necessary as Australian researchers tend to move between institutions.

Possible areas of development include:

- liaison with DEST RQF staff to develop information requirements
- development of RQF base "schemata" and "vocabularies" for the sector at large
- repository support for customised schemata and controlled vocabulary

2.2. Complex or Non-text Items

The RQF will necessarily want to assess complex digital items such as data sets, image banks, musical scores and performances, multimedia presentations, blogs, web sites etc. The repositories will need to assess their ability to hold or reference such works.

2.3. Work Flows

Repositories will need to map easy work flows to guide the submission of research items to their RQF collections. The alternative will be ingesting thousands of items on behalf of the researchers. The repository systems will need to enable self-submission whilst still ensuring crucial metadata is collected at the time of submission. "Smart" systems that leverage off existing information will be most effective (for example by integrating with personal details from the institution's HR and admin directories).

Other university administrators (from the research office or the research group itself) will surely want to check off the validity and completeness of the submitted items, and they will benefit from configurable workflows to suit their needs.

One would expect DEST to provide the workflow tools to assessors, however until the implementation is finalised we won't know if the institution will be expected to provide any specific functionality to the assessors.

Possible areas of development include:

- development of appropriate generic workflow descriptions for use in the sector at large
- repository support for customisable workflows

2.4. Group Management and Access Management

The repository system will need to know which groups of people have permission to submit or assess items in the RQF collections. The RQF model thus far includes a notion of funding being contingent on the outcomes of the assessment. This raises the stakes and means that submitters, assessors, and administrators will need to be identified and authorised with their appropriate permissions.

"Identification" means a login and password and these should be integrated into the university's existing authentication systems. External assessors will need to be identified through some other system (either a complex federated shibboleth system or a locally maintained group of custom logins and passwords).

Having identified the researcher, the next step is authorisation. The repository system will need to know which collections to "authorise" access to. This will mean essentially the maintenance of hundreds of constantly changing groups. If these groups are to be maintained within the repository, repository administrators will need to be able to delegate group management to administrators of the different research groups. Or if the research groups are defined elsewhere in the university's enterprise information system (an identity server, for example) then they need to be integrated with or uploaded into the repository's authorisation system. The latter is more elegant long term, but may not be feasible in the short term.

Complex access management systems can be developed, but they rely upon the existence of well maintained group membership, from which the individual's attributes may be inferred.

Possible areas of development include:

- description of generic access management requirements for RQF
- development of group management and authorisation systems for repositories
- development of external identity services to define the research communities involved in RQF

2.5. Communication and Reporting

Until the RQF implementation model is complete, we will not know if we are dealing with a model where the RQF harvests information directly from the repositories or whether the repository reports to the local research office who mediate the information. Until then, we don't know who the repository needs to communicate with. DEST and the institutions may have differing priorities here.

Either way one may confidently assume that either DEST or the research office (or both) will want to query the repository and make reports of the metadata in the repository. The OAI-PMH system (and other web enabled query systems) should be extended to facilitate this "harvesting" of information in the repository. The development of a standard data model (discussed above) will be all the more important in this context.

The RQF process itself is a one-in-five year process - a periodical report. It is likely that DEST and/ or the university research office will want to make interim reports about the research items held in the repository.

The repository may want to support real time queries, reports, and metadata harvesting.

Review of the JISC developed RAE communication models will give some direction to development in this area.

2.6. Liaison with the Research Office

In most universities the major RQF stakeholder will be the DVC (Research) and the Research Office. The repository service will usually report to a different part of the university and will be playing a support role to the Research Office in the area of information management for the RQF. Obviously the repository service will need to develop good working relationships with the Research Office. There will certainly be a need for repository software systems to integrate in some way with the research management software used locally by the Research Office. Moreover the repository service will need to analyse the needs of the local Research Office and modify data models, metadata, etc accordingly. Repository developers may want to go as far as providing "pre-packaged" integration modules for some of the more popular research management software.