

**Repositories –ERA 2009 HCA Trial  
For CAUL/CAIRSS Meeting (February 2010)**

*This is an overview of the key repository-related experiences arising in the 2009 ERA Trial for the Humanities and Creative Arts (HCA) disciplinary cluster.*

In 2009, the testing of research outputs (ROs) in institutions' repositories was undertaken in three stages:

- (1) Testing for the existing host (repository);
- (2) Testing for the sample and the full-submission urls (links); and
- (3) Testing for SEER access to the research output (RO)
  - a. Where the Research Output Digital Asset (RODA) is hosted by the institution;
  - b. Where the RODA is hosted and maintained by an organisation external to the eligible institution, such as a publisher, Google Scholar, the National Library, and others.

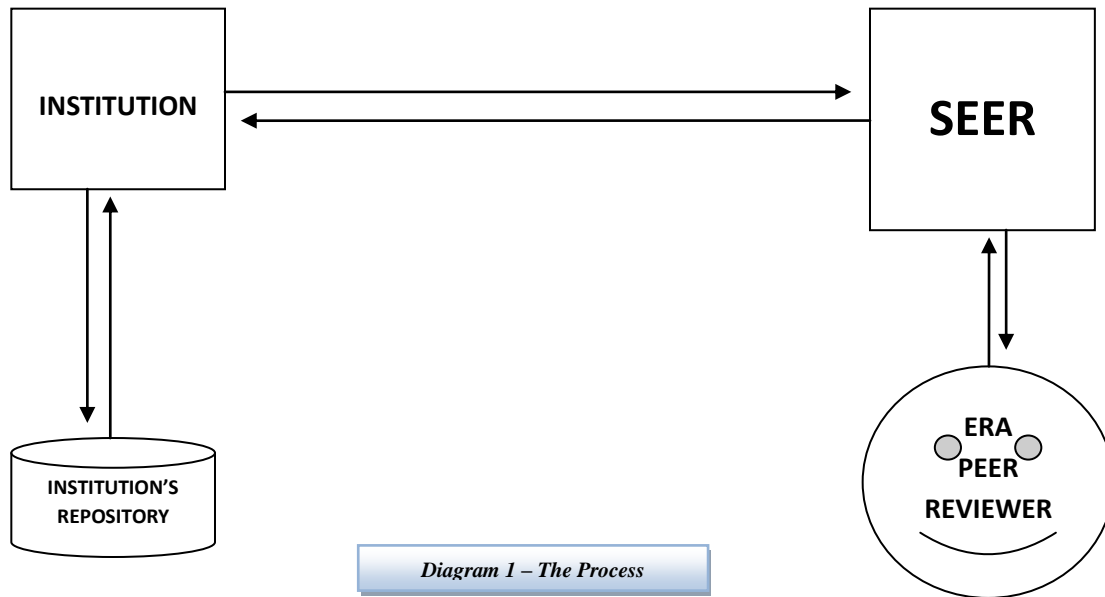
Stages (1) and (2) were largely the focus of repository testing in 2009, as the working model of SEER was still being fully developed and implemented. Thus, while the majority of repository-related issues concerning independent access to repository links and ROs were worked out early on, the critical question of access to these through the SEER interface (Stage 3) was an issue.

In particular, there were the complexities associated with:

- Data travelling through a number of points (nodes), and being repackaged for anonymity; and
- Maintaining the anonymity of (i) the ERA peer reviewer from the institution or other host organisation, and (ii) the repository data from the ERA peer reviewer.

These processes are best illustrated through a series of diagrams on the following pages.

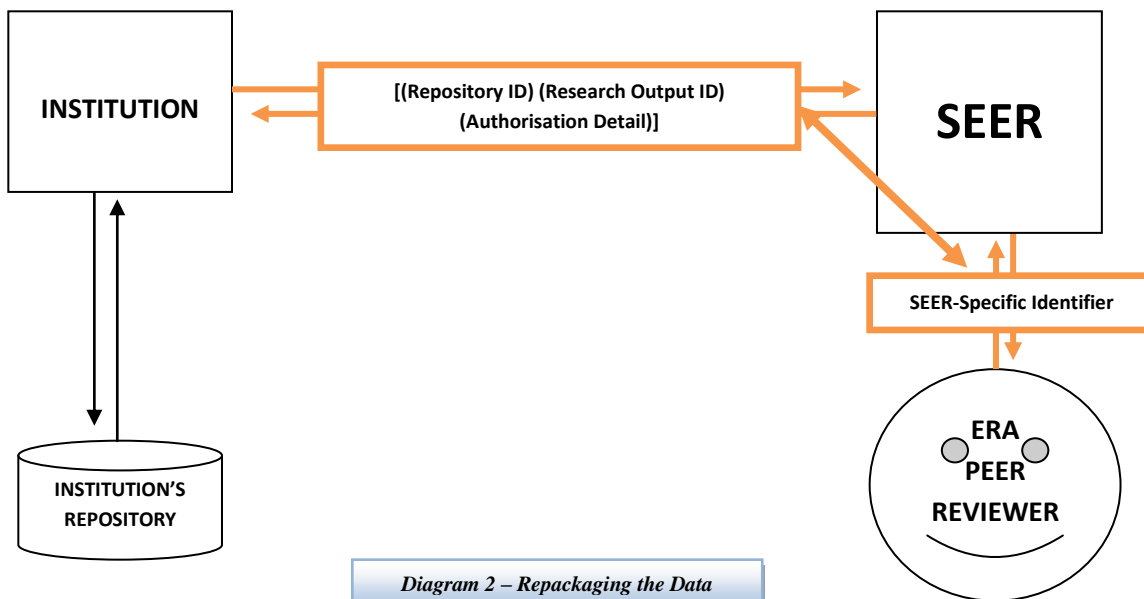
Diagram 1, below, illustrates how an RO travels virtually from an institution's repository through a number of nodes to the ERA peer reviewer and back again.



Each RO submitted by an institution for ERA peer review has a separate link into that institution's repository. These links contain two, sometimes three, components:<sup>1</sup>

[[Repository ID] (Research Output ID) (Authorisation Detail)]

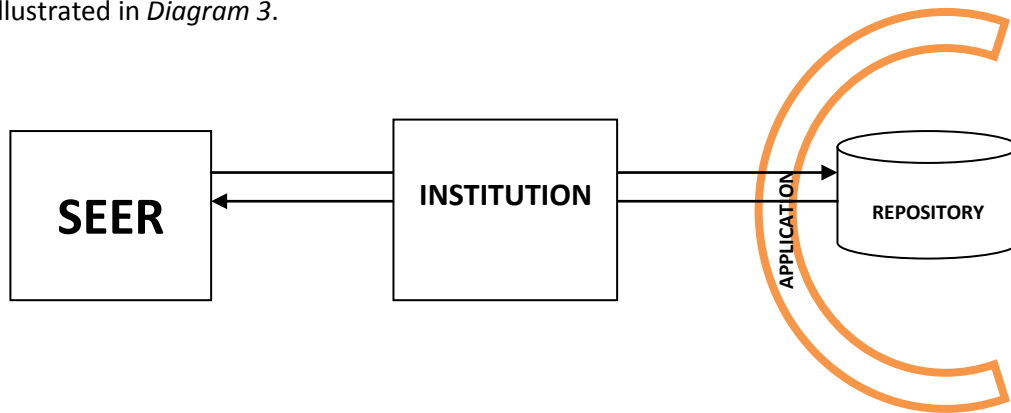
In order to maintain the anonymity of both the repository from the ERA peer reviewer and the ERA peer reviewer from the institution, SEER must repackage the components of the submitted link into a unique and SEER-specific identifier. *Diagram 2* illustrates the process of this repackaging of data.



<sup>1</sup> The third component, the 'Authorisation Detail', will be present if the institution or host organisation uses basic authentication.

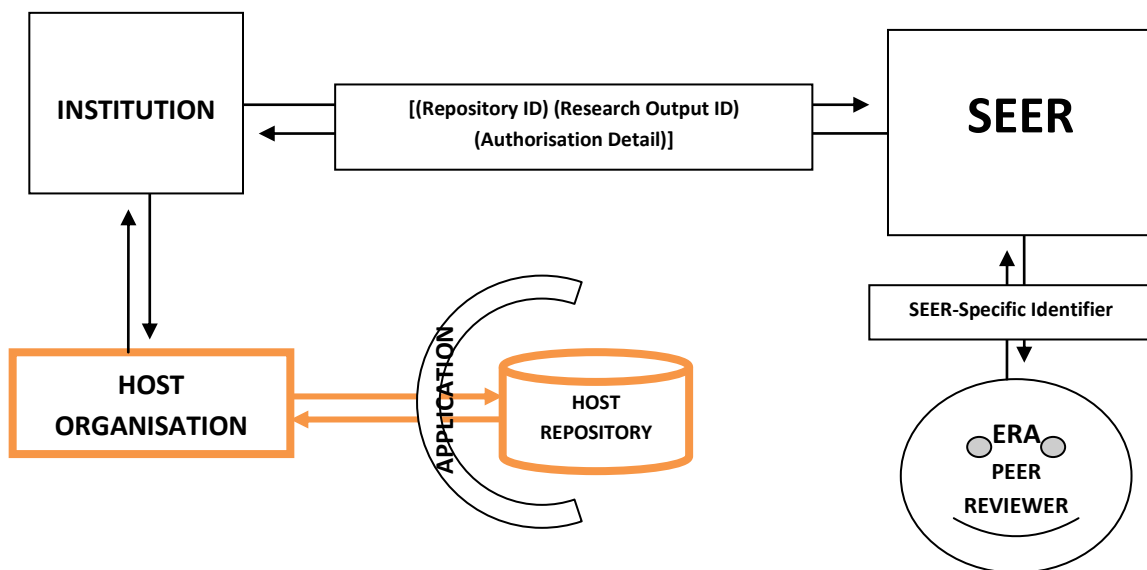
The repository issues in the ERA 2009 HCA Trial were also complicated by two other factors.

Firstly, there was an additional interaction point, an **application**, under which an institution's repository runs. This proved an additional wall between SEER's access to the institution's repository, as is illustrated in *Diagram 3*.



*Diagram 3 – A Repository's Application*

Secondly, the process was also complicated when an institution used **another organisation's repository**. This extra node that an RO needed to travel in order to reach its final destination (the ERA peer reviewer) provided a further breakdown point, as illustrated in *Diagram 4*.



*Diagram 4 – Using a Host Organisation's Repository*