WAGGING THE LONG TAIL OF RESEARCH DATA: IMPROVING THE VISIBILITY AND INTEROPERABILITY OF DISPARATE DATASETS

Objectives:
Universities and research institutions are becoming increasingly interested in collecting and providing access to datasets produced at their institution that do not fall within the scope of other discipline-based, or government repositories. However, there are numerous challenges. There is a huge range and diversity of datasets created in this context. In addition, there are significant variations in regional approaches. Some countries have invested in discipline-based repositories, while others are focusing on institutional data repositories. Furthermore many datasets do not adhere to appropriate standards. For example, the Science survey states, “Even within a single institution there are no standards for storing data, so each lab, or often each fellow, uses ad hoc approaches.”( Science 11 February 2011: Vol. 331 no. 6018 pp. 692-693 DOI: 10.1126/science.331.6018.692). How can we fit local set-ups, which are different from country to country, and disciplinary difference, into a global set of good practice advice?
The aim of this interest group is to develop a set of good practices for managing research data archived in the university context. The scope of the topic will be limited to the data generated in universities and research institutions and the role of institutional repositories and libraries as agents of the institutional data management.

Activities:
The interest group will begin its work in September 2013 and undertake the following activities:
• Map the current repository landscape: to categorize the types of repositories that do exist along a number of axes such as domain, open/closed, data formats, etc.; to have a understand the capacity of existing repositories to collect small datasets; and to identify gaps
• Define the scope of datasets/repositories that will be addressed in this project.
• Review data federation approaches that provide mechanisms for supporting discovery across the myriad of existing repositories (distinguish domain specific attributes from generalizable practices)
• Develop a number of use cases, based on a range of disciplinary practices and other approaches.
• Identify and publish good practices.
• Identify skills and competencies for those managing research data in university repositories.

URL: https://rd-alliance.org/internal-groups/long-tail-research-data-ig.html