Commentary: Whirlwind Tour of Digital Curation in the UK
- Karen Baker (CCE/PAL)

Two representatives from CCE and PAL LTER attended the 4th International Digital Curation Conference (DCC) in Edinburgh 1-3 December 2008 as part of the NSF/SES HSD Comparative Interoperability Project. The DCC (www.dcc.ac.uk) is a UK initiative to facilitate data-centric curation by providing an annual international conference, data policy, learning support, curation tools, training and a journal, the International Journal of Digital Curation (IJDC). Conference participants brought diverse perspectives from such fields as institutional data libraries, national data archives, computational software designers, large- and local- scale scientists, social scientists and others. Lynn Yarmey and I traveled to Scotland to present our paper "Data Stewardship: Environmental Data Curation and a Web-of-Repositories" as well as to gain some international perspective on data curation and infrastructure initiatives.

Within the U.S. as well as Europe, there are a variety of contemporary national and institutional efforts with respect to digital data preservation. Most frequently using open-source repository software – Eprints, DSpace, Fedora – national initiatives and institutional libraries are developing digital infrastructure and exchange capabilities with a focus on data preservation. Workshops on the first day of the conference covered the topics of the DCC Curation Lifecycle Model, a Data Audit Framework, and Repository Curation Service Environments. A trio of terms was central to the Lifecycle Workshop and to subsequent DCC discussions:

1) Data Curation – maintaining and adding value to a trusted body of digital information for current and future use

2) Digital Repository – applied to digital storage initiative such as institutional repositories, digital archives, or digital data libraries each with different functionality but having an organizational framework and a technical infrastructure

3) Curation Life Cycle Model – a model that incorporates curation and preservation stages including a) full lifecycle actions such as description and planning to b) sequential actions from creation or receipt through appraisal, ingestion, storage, access/use/reuse and transformation and c) occasional actions such as dispose, reappraise and migrate.

In our paper, we emphasized the importance conceptually of diverse repositories brought together eventually into a 'web of repositories' but also highlighted the notion of 'local repositories' as a critical but overlooked repository type. With the LTER network model of embedded data management in mind, we established a boundary in order to distinguish 'near' and 'remote' repositories. Repository characteristics were then explored starting with the characteristic difference in terms of a repository's 'distance-from-origin' from the data and the repository's focus, ie on field data or well-defined data objects. Within this framework, the LTER network represents 26 local repositories that are uniquely 'near' in terms of distance-from-origin. Each repository carries out local data curation with a deep involvement in creating datasets for immediate use as well as long-term reuse.

The visit afforded some additional special opportunities. On a walking tour of the narrow winds, we learned about 'gardyloo' and 'hud yer hand' as warning cries to proceed carefully. This brought to mind an early server given the name 'gardyloo'. We were able to meet with those involved with data management at the British Antarctic Survey, a PAL partner, and those at the Natural Environment Research Council (NERC) sponsored British Oceanographic Data Center (BODC). In addition, attendance at a series of talks commemorating the 25th anniversary of the University of Edinburgh Data Library provided further insight into past and present data efforts. Finally, there was a special conference dinner at the Edinburgh castle. With the strains of bagpipes welcoming us, we had a chance to view some distinctly non-digital artifacts – Scotland's crown jewels – where metadata was provided in the form of a docent's lively narrative that momentarily brought to life a wee bit of Scottish history.