Geophysical Research Abstracts, Vol. 9, 08458, 2007
SRef-ID: 1607-7962/gra/EGU2007-A-08458

Safeguarding the Citation Lifecycle for Global Geospatial Repositories

R. Bose (1), G. McGarva (2)

(1) School of Informatics, University of Edinburgh, United Kingdom, (2) EDINA/Data Library Services, University of Edinburgh, United Kingdom

During the next decade more geoscientific institutions around the globe will house their geospatial data collections in geodatabases (using extended relational database systems like Oracle Spatial or varieties of the ESRI Geodatabase) or geospatial repositories (based on emerging concepts of digital object containers as in, for example, DSpace and Fedora repositories). These data collections will form the basis of collaboration in the geosciences, and will be at least partially accessible to international collaborators through distributed queries, web service requests, and the like. As a consequence, we can expect that: (1) various subsets of the contents of these databases and repositories will be cited by scientific studies; and (2) international researchers will both want and need to resolve these citations over time. To keep the scientific record intact, and ensure that geospatial data cited today remains available to researchers in the future, the emerging principles of digital data curation (long-term preservation and access) must be emphasized among geoscientists and scientific information specialists and managers. Data curation principles and practices include designing archives that facilitate version control and provenance retrieval, which are ultimately related to creating and resolving citations. Our work investigates the design of systems that incorporate curation principles to help safeguard the citation lifecycle for geospatial data collections.