

Spring, 2006



01001100 01010100 01000101 01010010
LTER DataBits
 Information Management Newsletter of
 The Long Term Ecological Research Network
 01001100 01010100 01000101 01010010

From Databases to Dataspaces: Opening up Data Processes

- *Karen Baker (PAL/CCE)*

Franklin, M., A. Halevy, and D. Maier, 2005. From Databases to Dataspaces: A New Abstraction for Information Management. SIGMOD Record 34(4): 27-33.

In the complex work of bridging data collection and data federation, our conceptual understandings grow regarding LTER research site data. These data efforts, informed and guided by local scientific needs and conventions, entail alignment of technological approaches and development of community nomenclature, standards, and dictionaries. Amidst such activities, this paper 'From Databases to Dataspaces' offers 'a new abstraction' that seems to escape the confines of a traditional data box model.

The authors are part of a group that has met over the last years to consider in depth the development, functionality, and use of data base management systems. This recent work opens up the data landscape conceptually - from databases to dataspaces. The 'dataspaces' approach is presented as both a new agenda and an architecture that allows for multiple ways of solving issues and framing questions of information management. In addition, development is recognized to take place over multiple timeframes: "One of the key properties of dataspaces is that semantic integration evolves over time and only where needed. The most scarce resource available for semantic integration is human attention." The dataspace concept umbrellas explicitly a continuum of organizational and semantic arrangements that handle diverse data types, states, and approaches. The paper recognizes the multiple facets of work involved: "Dataspaces are not a data integration approach; rather, they are more of a data co-existence approach". In the paper, there is discussion of dataspace requirements, components, and research challenges.

Perhaps this paper struck a chord because it seems to resonate with one of my first impressions of LTER, that is, a community recognizing the value of a good number of distinctly unique LTER sites working in loose proximity. For LTER, years of joint projects addressing local, regional, and cross-site science have contributed to a thick infrastructure. This infrastructure includes a shared sensitivity toward respect for local diversity and trust that new knowledge as well as research challenges arise from grappling with heterogeneity. As we seek tighter collaborative configurations in both ecological science and informatics endeavors, the dataspaces concept highlights the notion of a development process over time of data and information management, of theory and practice.