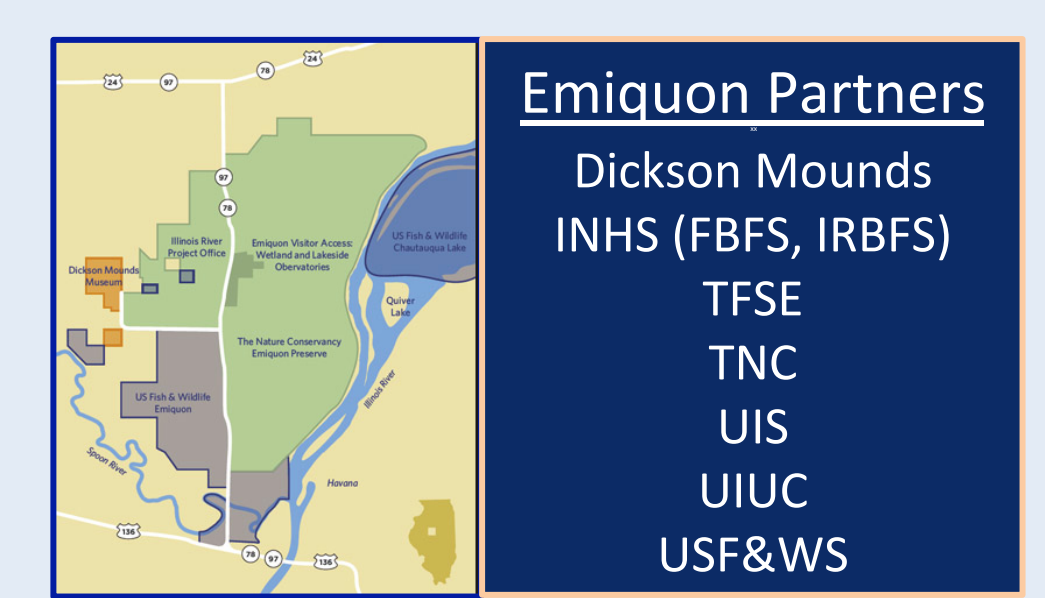


Planning for Data Management at Emiquon

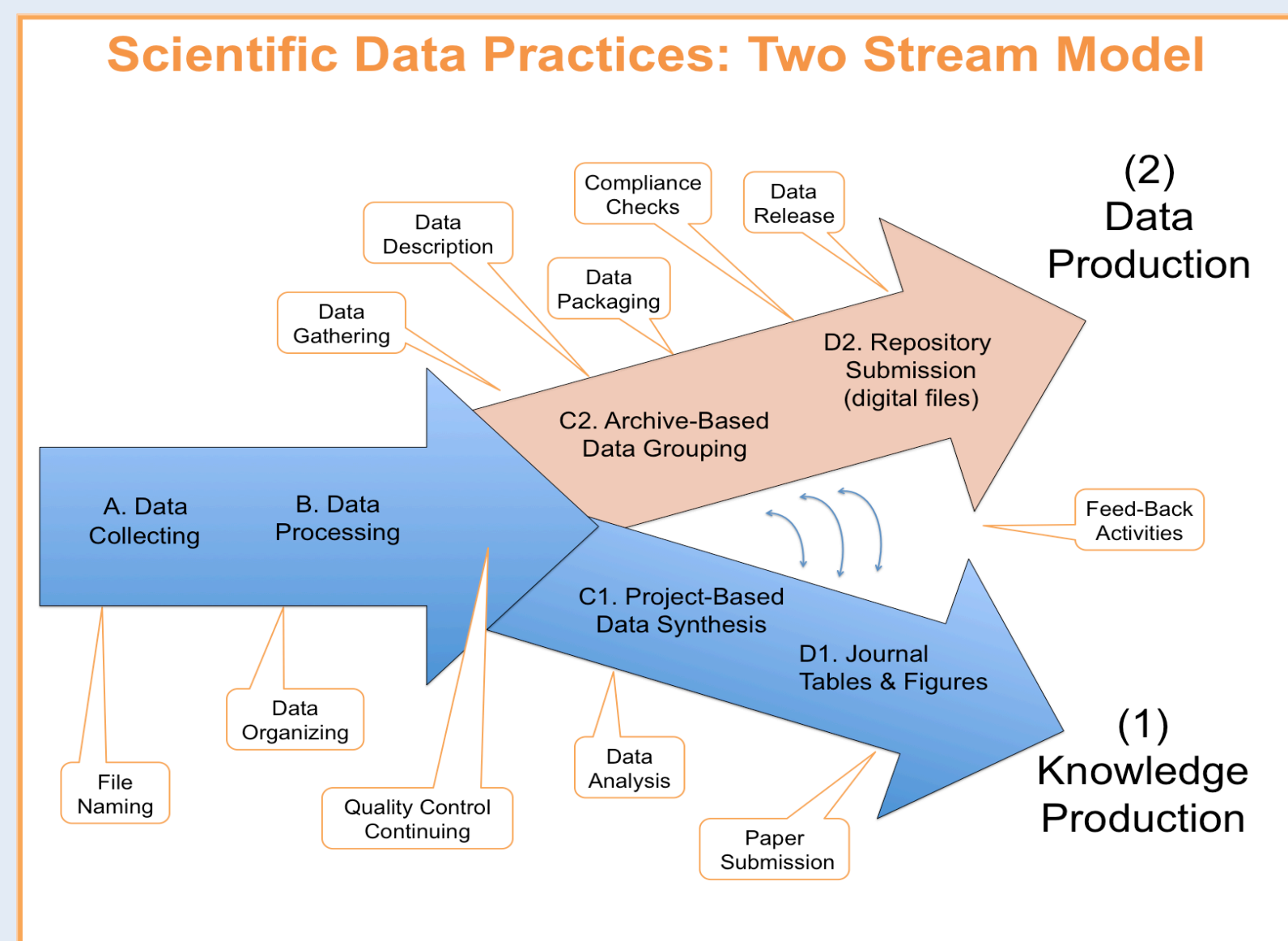
Karen S. Baker, University of Illinois Urbana-Champaign, karenbaker@gmail.com



Introduction

Current practices and configurations are not enough to realize the vision of data as open, accessible, and reuseable resources. While data are central to the production of scientific publications, additional work is required to make field-oriented, environmental science data available for shared use.

New Data Work



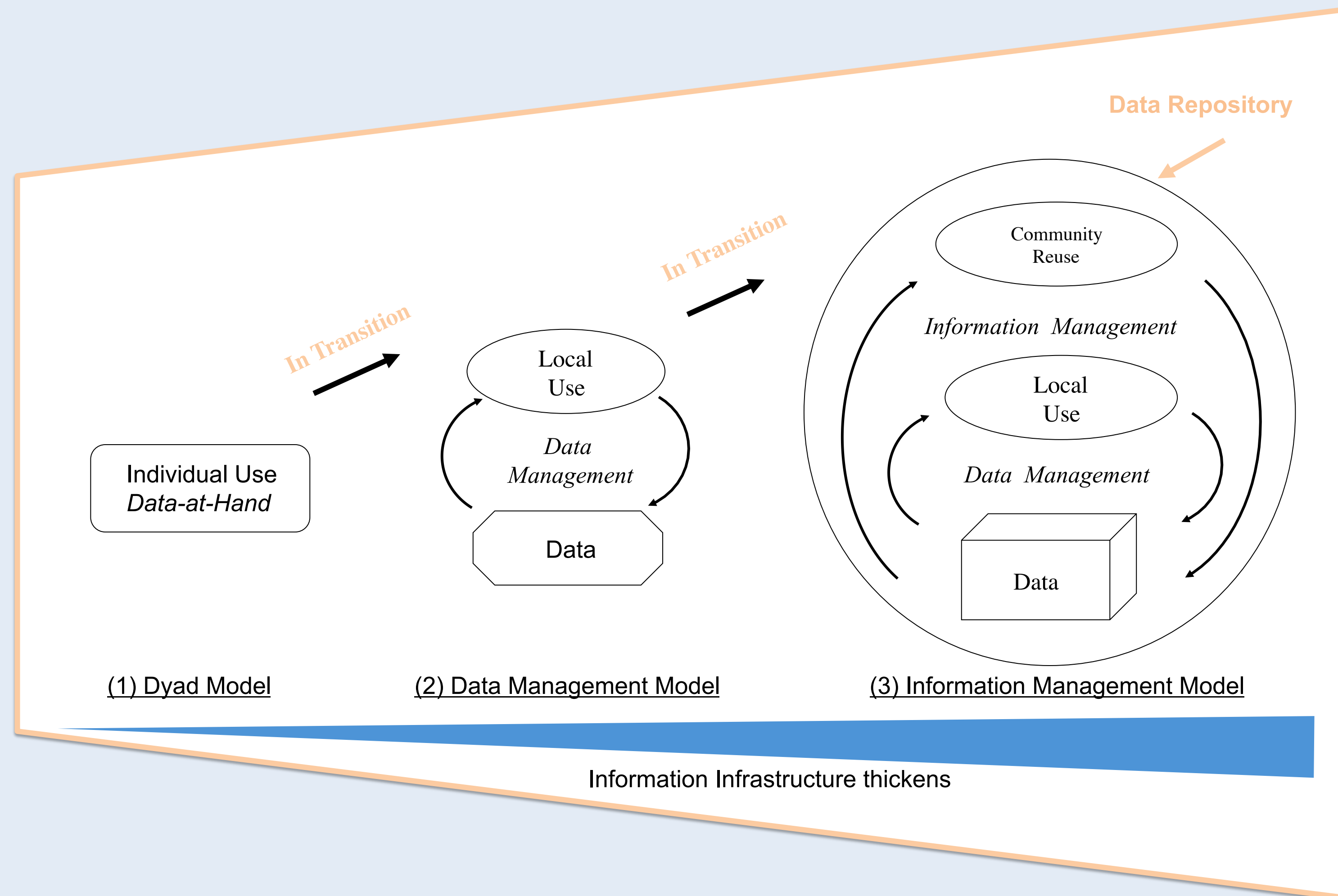
(1) Knowledge Production

Data for internal use (shown as blue arrows) are subject to a complex mix of processing, analysis, integration, and presentation strategies with the final form optimized for publication of papers.

(2) Data production

Data for public reuse (shown as red arrow) are prepared with a more standardized set of procedures to create well-described, parameter-based sets of data for release to a data repository and for reuse by others.

Developing Infrastructure



Among the challenges that arise with data aggregation and data sharing are the transitions required from one-time informal exchanges to a more formal data production process. Change depends upon development of new data activities, expertise focusing on management of data, and growth of situated information infrastructure. A data management trajectory is shown above developing in three phases: (1) an isolated research laboratory with perhaps a researcher and a research assistant, (2) a data management model where data are aggregated and shared with an emphasis on local use, and (3) an information management model where data reside within a data repository that supports data discovery and delivery.

-Baker and Millerand, 2010. Infrastructuring ecology: Challenges in achieving data sharing. In Collaboration in the New Life Sciences, Ashgate.

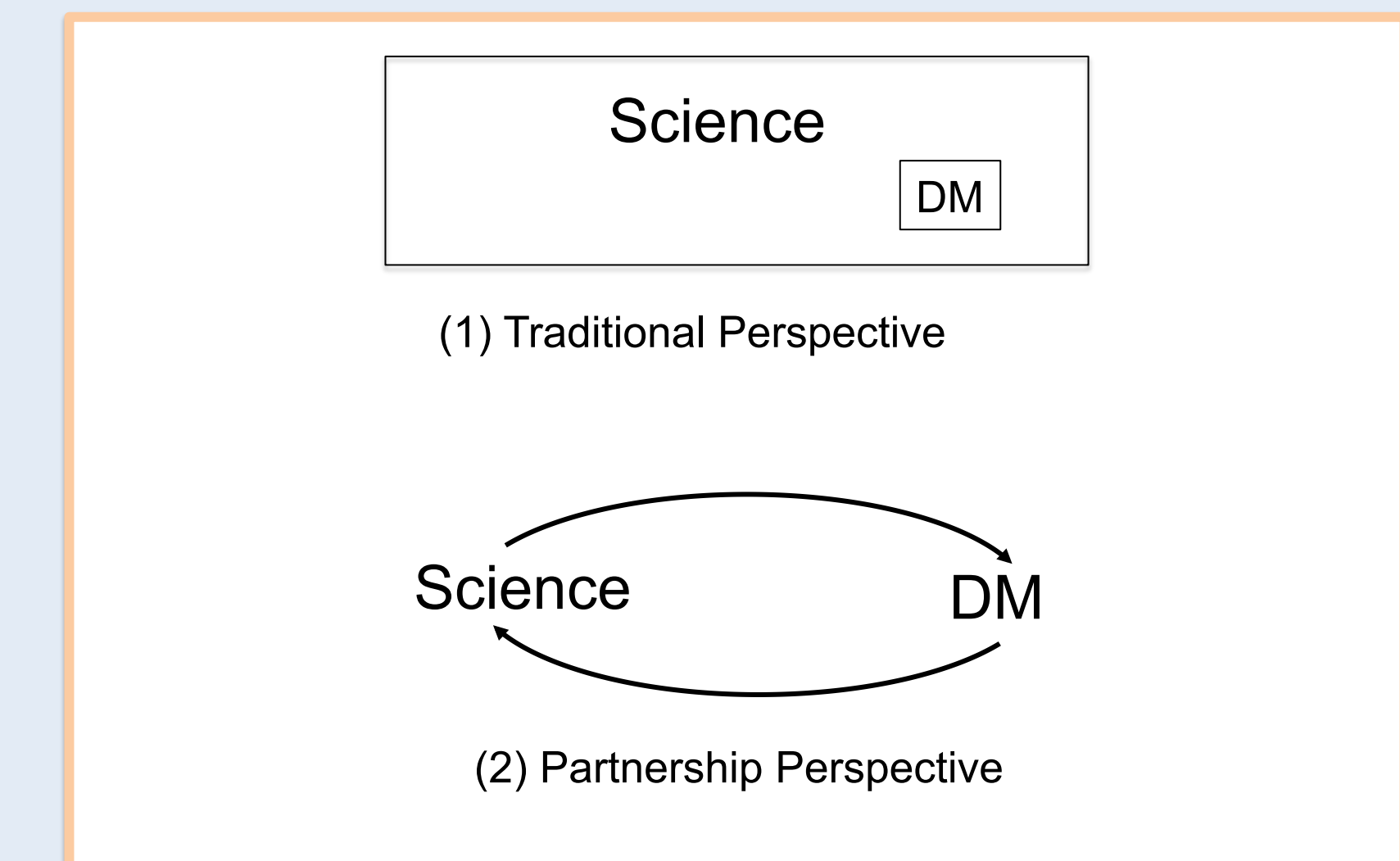
Acknowledgments

This work is supported by the Institute of Museum and Library Services (IMLS) through funding for the Data Curation Education at Research Centers (DCERC; IMLS Award# RE-02-10-0004-10) and the Site-Based Data Curation Project at Yellowstone National Park (SBDC; IMLS Award# LG-06-12-0706-12).

Changing Data Practices

- New work is required for sharing research data
- Aggregation of data from multiple groups initiates transitions in data arrangements
- Infrastructure growth is key to collaborative data efforts
- Transitions in data work include development of data management for data use locally and information management for data reuse by additional designated communities

Developing Partnership



(1) Traditional Perspective

Scientific research has embedded within it the processing, analysis, and synthesis of data.

(2) Partnership Perspective

Scientific research partners with data management for organizing, sharing and preserving of data.