Sustainable Data Curation Infrastructure

Indo-US Workshop
March 24, 2009
Sayeed Choudhury

Sustainability

"What it means is that the practice or process can't go on indefinitely because it is destroying the very conditions on which it depends."

> – Michael Pollan (author on sustainable food practices)

Data Conservancy

- Pending award through the National Science Foundation DataNet Program

- The Data Conservancy (DC) embraces a shared vision: data curation is not an end, but rather a means to collect, organize, validate, preserve and make available data to address the grand research challenges that face society

- DC's sustainability plans encompass both technological and economic dimensions

JOHNS HOPKINS
UNIVERSITY

Technological Sustainability

- Technical details described in workshop paper

- From a sustainability perspective, the important point is to build upon existing systems and prototype development efforts (e.g., Fedora Commons, Sakai, and Microsoft Research's research output repository and Research Information Centre), *which are based on technical requirements*

Economic Sustainability

- Alignment of existing institutional priorities with Data Conservancy goals

- Johns Hopkins University Sheridan Libraries experience with digital astronomy and digital manuscript data (economies of scale and economies of scope)

- Data Conservancy partner institutions' efforts related to information science research agenda, *which will help define scientific requirements*

JOHNS HOPKINS
UNIVERSITY

Blue Ribbon Task Force on Sustainable Digital Preservation and Access
http://brtf.sdsc.edu

- Conduct an analysis of previous and current models for sustainable digital preservation, and identify current best practices among existing collections, repositories and analogous enterprises

- Develop a set of economically viable recommendations to catalyze the development of reliable strategies for the preservation of digital information

- Provide a research agenda to organize and motivate future work in the specific area of economic sustainability of digital information

Blue Ribbon Task Force on Sustainable Digital Preservation and Access
http://brtf.sdsc.edu

- First year report focused on the "positive" (or "what is" or the state of the art) environmental scan

- Second year report will focus on "normative" (or "what ought to be" or recommendations)

- Broadly speaking, there has been accounting of costs, some consideration of revenues, but no examination of full economic model

JOHNS HOPKINS
U N I V E R S I T Y

Final Thoughts

The most important point to stress is that the sustainability team *will introduce business requirements* at the *beginning* of DC's data infrastructure development effort

Technical Requirements ⟷ Cost Models

Scientific Requirements ⟷ Revenue Models

Business Requirements ⟷ Economic Models

JOHNS HOPKINS
UNIVERSITY

Acknowledgements

DataNet award for "The Data Conservancy"

NLG grant award LG0606018206

sayeed@jhu.edu   Indo-US Workshop

JOHNS HOPKINS
U N I V E R S I T Y

Thank You!

Sayeed Choudhury
Sheridan Libraries
Johns Hopkins University
Baltimore, MD 21218
USA

+1-410-516-4930
sayeed@jhu.edu

JOHNS HOPKINS
U N I V E R S I T Y