The Data Capsule for Non-Consumptive Research
Beth Plale, Atul Prakash, Geoffrey Fox, Robert H. McDonald

A Proposal to the Alfred P. Sloan Foundation

HTRC Kickoff Reception
DH2011 June 20, 2011
Non-Consumptive Research (NCR)

No action or set of actions on the part of HathiTrust Research Center users, either acting alone or in cooperation with other users over the duration of one or multiple sessions can result in sufficient information gathered from the HathiTrust collection to reassemble pages from the collection.
Data Capsule Framework (DCF)

The data capsule framework is a software framework utilizing software and hardware infrastructure at the IU location of the HathiTrust Research Center and utilizing software and hardware at IU that are part of the NSF funded FutureGrid project. This framework will allow researcher driven application use within a virtualized secure data intensive computing environment in order to provide general research use of the HathiTrust Digital Library Corpus.
In order to provide access to copyrighted content for research purposes, the researcher needs flexible computing resources in a controlled environment.

HathiTrust Digital Library Content
- Access to HT open content indices
- Access to HT copyrighted indices
- Auditable Secure Mechanisms for legal mandated MOU based and fair-use compliance

HTRC Managed Data Intensive Compute Resources

Secure Data Capsule Researcher Access

Researcher Driven Applications for Use as Services within the Data Capsule
- Can we provide a services framework for researcher applications to run within the secure data capsule compute resources?
More on NCR and DCF

Beth Plale
Director Data to Insight Center/Associate Professor of Computer Science
Indiana University
plale@cs.indiana.edu

Atul Prakash
Professor Electrical Engineering and Computers Science
University of Michigan
aprakash@eecs.umich.edu

Geoffrey Fox
Director Digital Science Center/Professor of Computer Science and Informatics
Indiana University
gcf@indiana.edu

Robert H. McDonald,
Associate Dean of Libraries/Associate Director Data to Insight Center
Indiana University
robert@indiana.edu