National Digital Information Infrastructure and Preservation Program (NDIIPP)
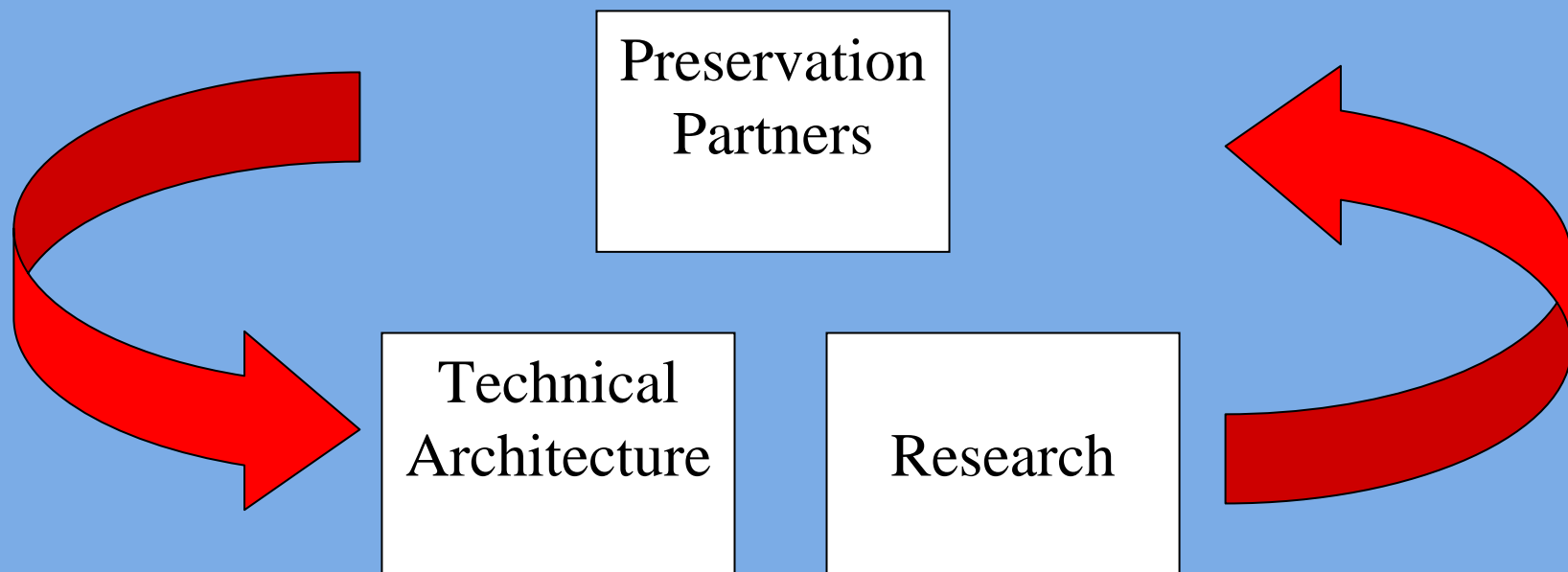
DLF Forum

November 2005

NDIIPP Elements

Preservation Partners

Technical Architecture

Research

Lead Institutions for the 8 Partnerships

- California Digital Library: Web political content
- Emory University: Southern digital culture
- UC Santa Barbara: Geospatial Data
- NC State University: Geospatial Data
- U of Maryland: Dot.com business records
- U of Michigan: Social science data
- U of Illinois at Urbana-Champaign: State government publications, among other content
- WNET/PBS: Digital television

Goals for the Partnerships

- Identify/select/collect content; communicate strategies for doing so
- Probe intellectual property issues
- Collaborate broadly in developing a shared technical architecture
- Study economic sustainability
- Identify and share best practices
- Learn how to build and incrementally improve a preservation network
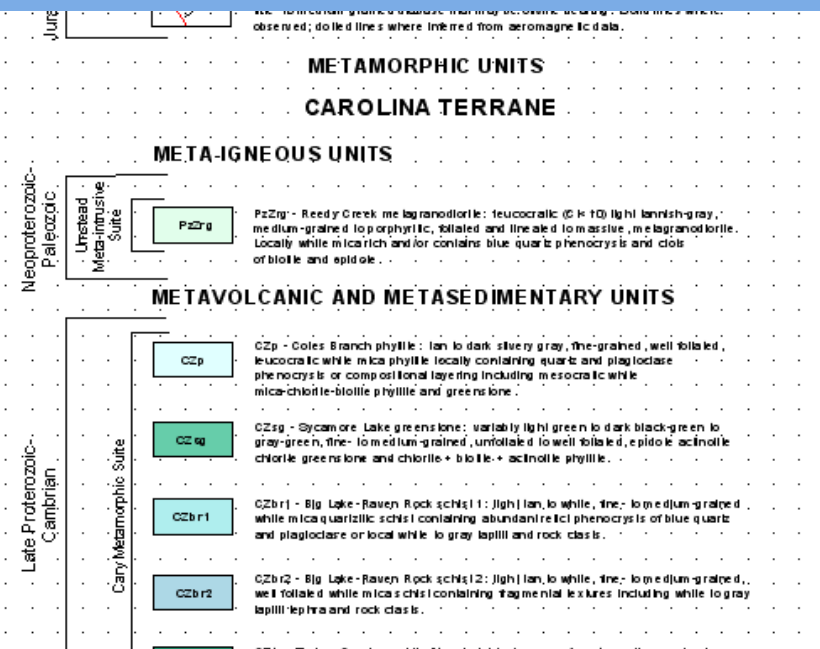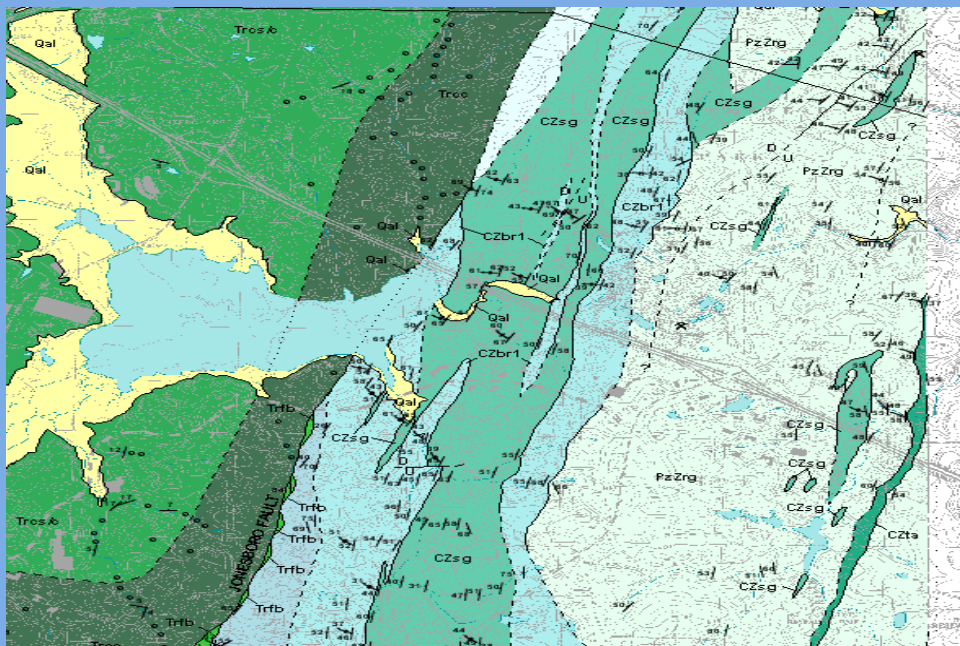
Network of Networks

- LC pleased to see interconnections between partners, other networks
- Some examples:
  - Association of Research Libraries
  - Digital Library Federation
  - Coalition for Networked Information
  - UK Joint Information Systems Committee

Broad Categories of Challenges

- Technical

- Legal

- Social

- Economic

- Handling complex objects in a repository
  - Multi-format datasets, projects vs. individual datasets
- Preserving spatial databases
  - The whole is more than the sum of the parts
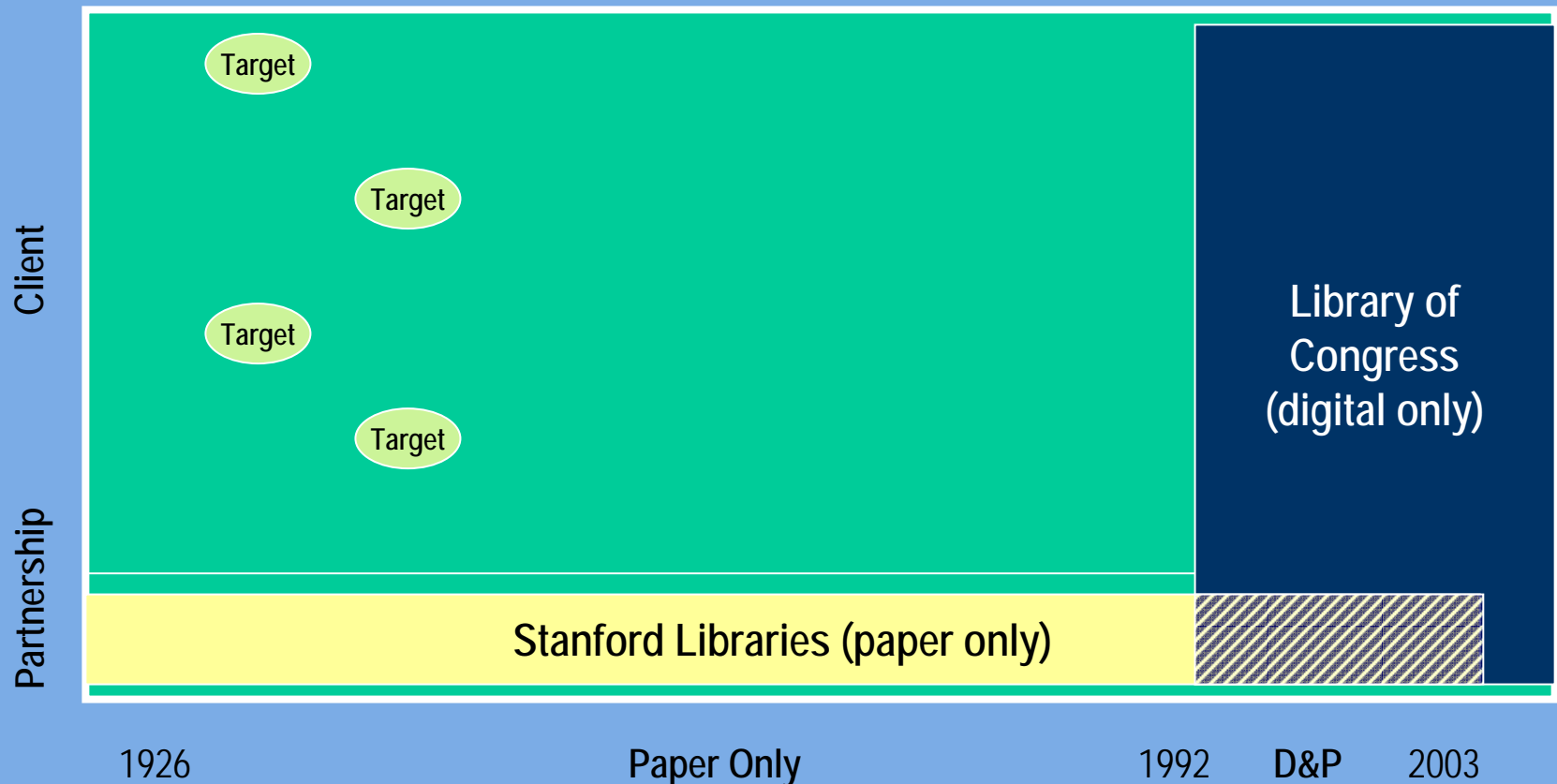- Managing time-versioned content

Emerging Issues

- Natural tension between collaboration and tradition of institution-specific approaches
- Much interest in joint infrastructure, but developing shared services will take time
- Existing preservation methods are being stretched
- Balancing a focus on project goals while also considering lots of interesting ideas
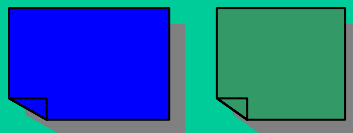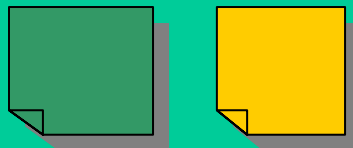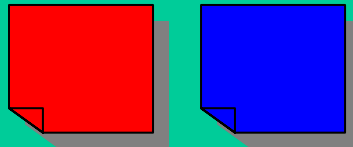
Brobeck Archive: Records Overview

Client

Partnership

Target

Target

Target

Target

Library of Congress (digital only)

Stanford Libraries (paper only)

1926   Paper Only   1992   D&P   2003

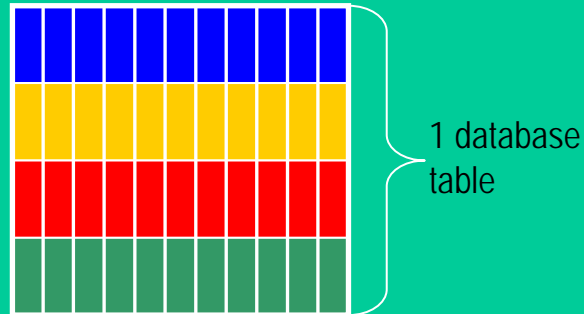# Brobeck Archive: Records Overview

Managed Documents (~3.48 million)

Word & Word Perfect (~3.34 million)
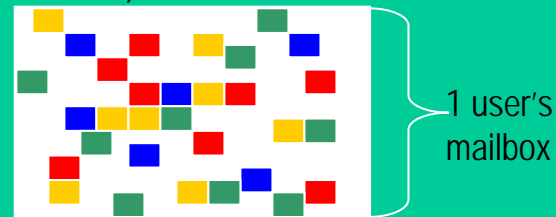
Excel (~54,000)

TIFF, PDF, PowerPoint, etc. (~90,000)
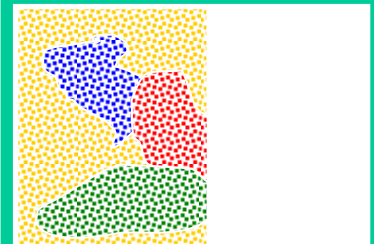
Relational Databases (~7)

1 database table

Brobeck Clients

- Client A
- Client B
- Client C
- Client D

Microsoft Outlook Data (~1,700 users; ~363GB)

1 user's mailbox

Network Share Drives

Closed Archive Classifications

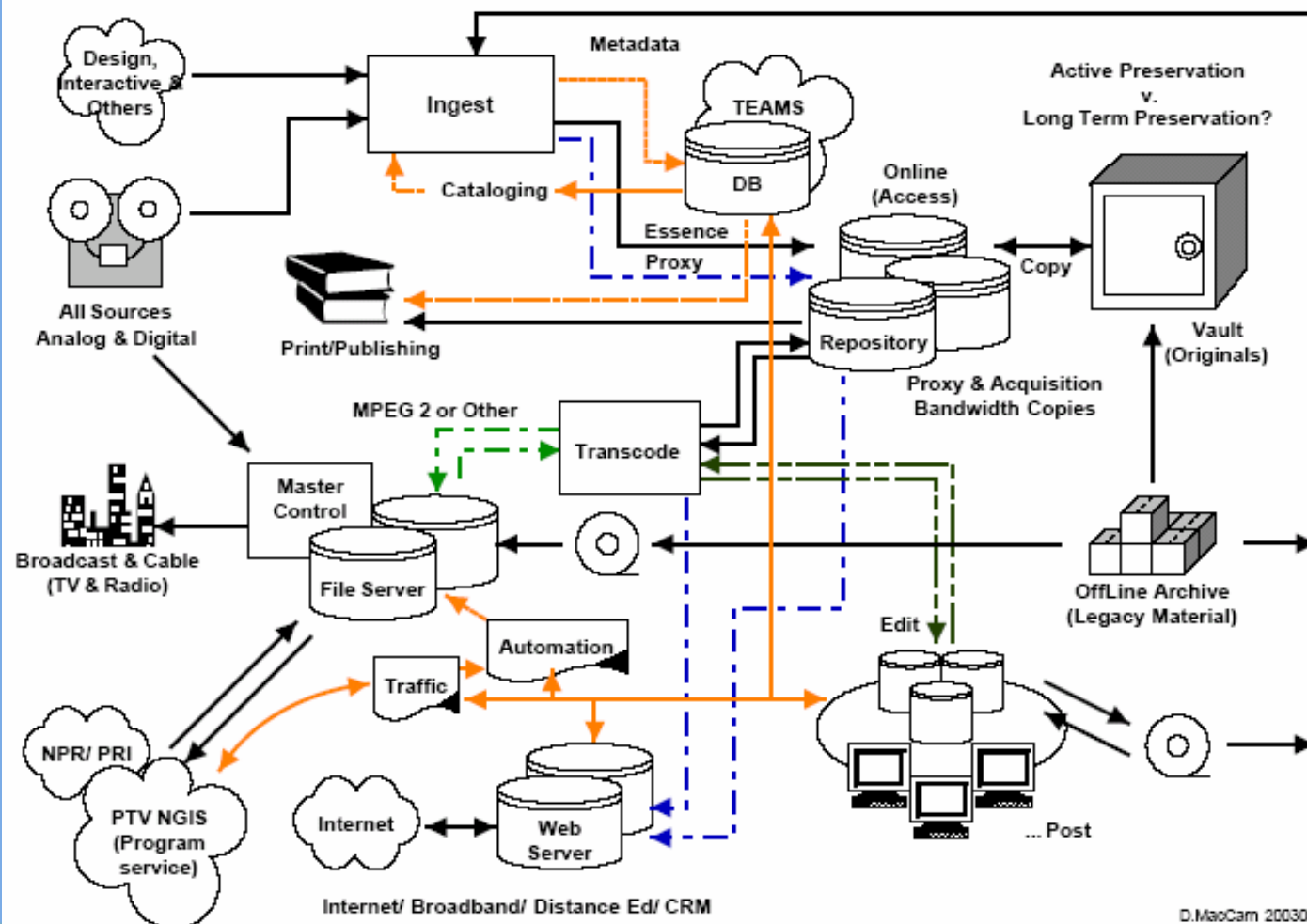|	Currently accessible under prevailing archival and legal rules	Currently confidential; potentially accessible in the future under prevailing archival and legal rules	Never accessible under prevailing archival and legal rules
Documents	Court-approved partnership documents Consenting clients' documents	Non-consenting clients' documents Non-court-approved partnership documents	Objecting clients' documents All personal clients' documents Sensitive partnership documents
Databases	Database rows or e-mails relating to consenting clients or court-approved partnership	All databases	None
E-Mail		All e-mail	None
Network Share Drives	None	All network share drives	None

Preserving Digital Public Television
What will this project do?

- Identify our "at-risk" *digital* holdings (we are not allowed to migrate any analog materials)

- Plan a long term preservation repository for America's digital public television programs

- Adopt technical standards and selection criteria for digital programs

- Design a test bed to develop operations for file transfers and a model repository

- Look at the finances of such a facility
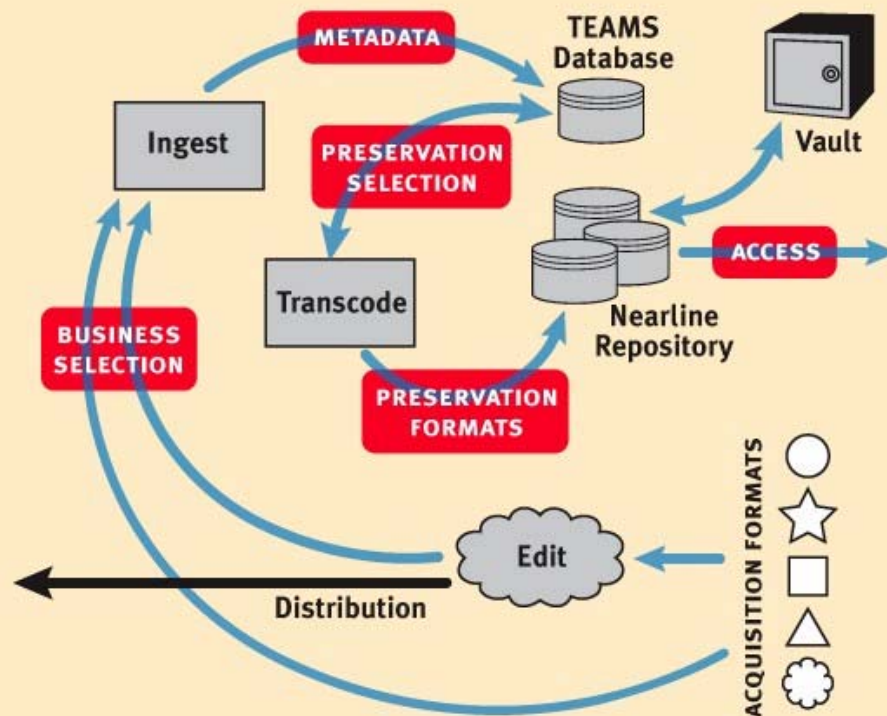
Digital Asset Management Unified Field Theory

Preserving Digital Public Television
PRESERVATION PATH

Learning and Incremental Development

- We need to get broad feedback and continue learning

- Still no "silver bullet" solution to digital preservation

- NDIIPP considering all viable approaches, working toward gradual development of decentralized, interoperable architecture

- Our partners are modeling this approach