DPN
THE DIGITAL PRESERVATION NETWORK

David Minor
UC San Diego Library
Chronopolis Preservation Network
What is DPN?
DPN = Digital Preservation Network, pronounced “Deepen.”
What is DPN?
One repository to rule them all!
One repository to rule them all!
DPN!
What Does DPN Do?

1. Establishes a network of heterogeneous, interoperable, trustworthy preservation repositories (Nodes)
2. Replicates content across the network, to multiple nodes
3. Enables restoration of preserved content to any node in the event of data loss, corruption or disaster
4. Ensures the ongoing preservation of digital information from depositors in the event of dissolution or divestment of depositors or an individual repository
5. Provides the option to (technically and legally) "brighten content" preserved in the network over time
DPN Benefits

1. Resilience
2. Succession
3. Economies of scale
4. Efficiency
5. Extensibility
6. Security
Critical Assumptions & Definitions

• All content enters DPN by deposit into one of the DPN Nodes.

• DPN Nodes can provide two general types of service:
  • “First Node” services: the point of entry for content;
  • “Replicating Node” services: holding dark copies of content in the DPN network.

• DPN Members will work directly with a First Node to negotiate contracts and determine service levels.

• Service levels and contracts will reflect “standard” DPN services; they may also reflect the First Node’s unique offerings in terms of access, hosting or other services.
Critical Assumptions & Definitions

• Content in Replicating Nodes will be held “dark”, and inaccessible except for preservation actions.

• DPN shall redistribute preserved content as Nodes enter and leave the Network, ensuring continuity of preservation services over time.

• Depositors, First Nodes and their designated communities will collaborate to ensure that the information contents of DPN deposits are accessible for reuse in the future, using the appropriate (and evolving) community standards for any given set of content.
Initial DPN technical partners

Initial DPN launch will feature five nodes:
• Academic Preservation Trust (APTrust)
• Chronopolis
• HathiTrust
• Stanford Digital Repository (SDR)
• University of Texas Data Repository (UTDR)

And a participating partner:
• DuraSpace
Infrastructure Components

- *Institutional Archive*
- *Federated Messaging (AMQP)*
- *Distributed Registry*
- *Transfer Mechanisms*
- *BagIt Content Packaging*
- *Security and Encryption*
Scenario 1: Ingest & Replication

1. Deposit

DPN Member

First Node

Replicating Node

2. Replicate

Replicating Node

2. Replicate
Scenario 2: Restoration of Content

1. Audit
2. (Red Cross)
3. Audit
4. Retrieve

Replicating Node

DPN Member

First Node

Replicating Node
Scenario 3: First Node Cessation

1. Fail
2. Brighten
3. Retrieve
Scenario 4: Successioning

1. DPN Member

2. Brighten

3. Retrieve

Future Member
So, where are we now?
60ish DPN members

<table>
<thead>
<tr>
<th>Member Listing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arizona State University</td>
</tr>
<tr>
<td>Brigham Young University</td>
</tr>
<tr>
<td>Brown University</td>
</tr>
<tr>
<td>California Institute of Technology</td>
</tr>
<tr>
<td>Columbia University</td>
</tr>
<tr>
<td>Cornell University</td>
</tr>
<tr>
<td>Dartmouth College</td>
</tr>
<tr>
<td>Duke University</td>
</tr>
<tr>
<td>Emory University</td>
</tr>
<tr>
<td>Harvard University</td>
</tr>
<tr>
<td>Indiana University</td>
</tr>
<tr>
<td>Iowa State University</td>
</tr>
<tr>
<td>Johns Hopkins University</td>
</tr>
<tr>
<td>Kansas State University</td>
</tr>
<tr>
<td>Massachusetts Institute of Technology</td>
</tr>
<tr>
<td>Michigan State University</td>
</tr>
<tr>
<td>University of Tennessee</td>
</tr>
<tr>
<td>University of Virginia</td>
</tr>
<tr>
<td>Utah State University</td>
</tr>
<tr>
<td>Yale University</td>
</tr>
<tr>
<td>Yale University</td>
</tr>
</tbody>
</table>
Technical Deliverables

☑ High level architecture & technical service design
☑ Registry design
☑ Packaging design
☑ Messaging development
☑ Date transfer testing
☐ Auditing (fixity) development
☐ Integration with local repositories
☐ Brightening tests & demonstrations
☐ DPN operational functions
   (reporting, billing, agreements, decryption, business audit, etc.)
☐ Scaling local infrastructure and operations
General timeline

*If you write this down, use a pencil.*

- Entering pilot phase with “real data”
- Heading towards soft rollout in fall 2015
- Full production in 2016
Thanks!

For more information:

http://www.dpn.org

inquiry@dpn.org