The Challenges of Building a Digital Preservation System

July 2011

Ido Peled

Ex Libris Rosetta Product Manager

ExLibris
The bridge to knowledge

Copyright Statement

ExLibris

What does *Google Videos* and *Blueberry Muffin* have in common ?
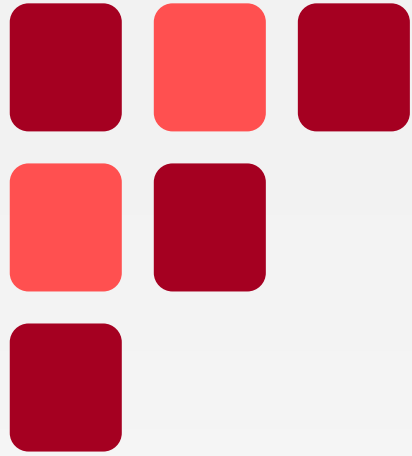
Why Digital Preservation

What is at Stake ?

"**Information is the oxygen of the modern age. It seeps through the walls topped by barbed wire, it wafts across the electrified borders**"

Ronald Reagan

Who Should Be Concerned ?

"A popular government without popular information or the means of acquiring it, is but a prologue to a farce, or a tragedy, or perhaps both"

James Madison

International Libraries	University Library	Researchers	Cultural Heritage Organizations
Media Companies	State & Local Governments	Federal Agencies	Specialist Archives
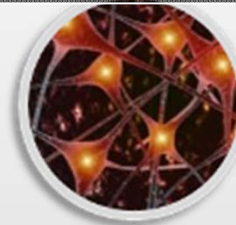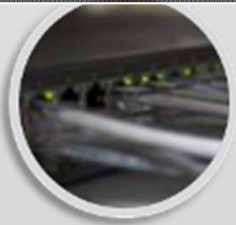
Challenges of Digital Preservation

Building a Digital Preservation System
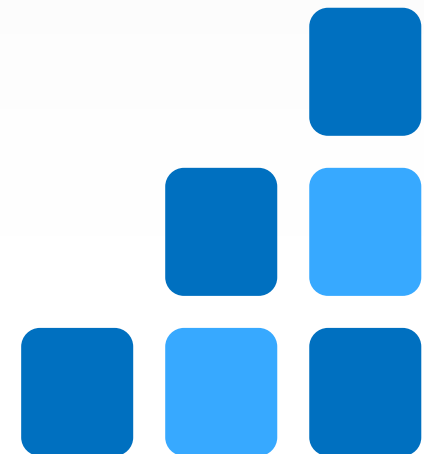
Challenge I – Digital Preservation Definition

ExLibris

Challenge I – Digital Preservation Definition

"**Digital preservation combines policies, strategies and actions to ensure access to reformatted and born digital content regardless of the challenges of media failure and technological change.**

The goal of digital preservation is the accurate rendering of authenticated content over time"

ALA (American Library Association)

Challenge I – Digital Preservation Definition

- Digital preservation is the management of digital information over time. Preservation of digital information is widely considered to require more constant and ongoing attention than preservation of other media. …
  en.wikipedia.org/wiki/Digital_preservation

- The long-term maintenance and upgrade of digital files on digital storage medium
  hosting.collectionsaustralia.net/capture/course/glossary.html

- The process of ensuring that a digital object is accessible over the long term.
  library.manoa.hawaii.edu/scholarspace/glossary.html

- Refers to the series of managed activities necessary to ensure continued access to digital materials for as long as necessary.
  www.digitaltraditions.net/html/Glossary.cfm

Challenge I – Digital Preservation Definition

Active Preservation

Digital
Preservation
Components

Archiving

Collection

ExLibris
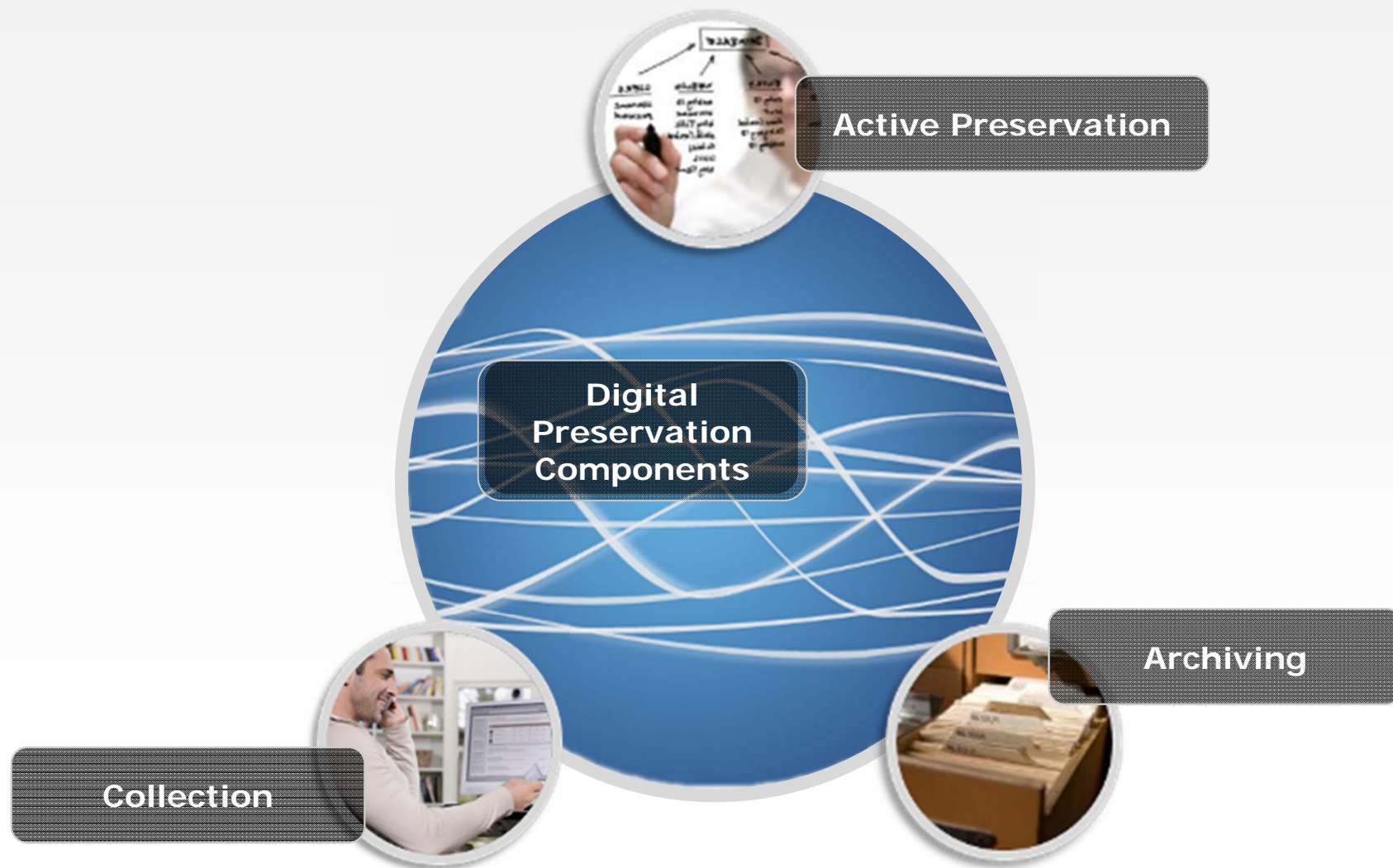
Challenge II – Data Model

- Support structure and metadata for long term digital preservation

- Variety of formats

- Variety of institution types

- Conform to standards and best practices

ExLibris

Challenge II – Data Model
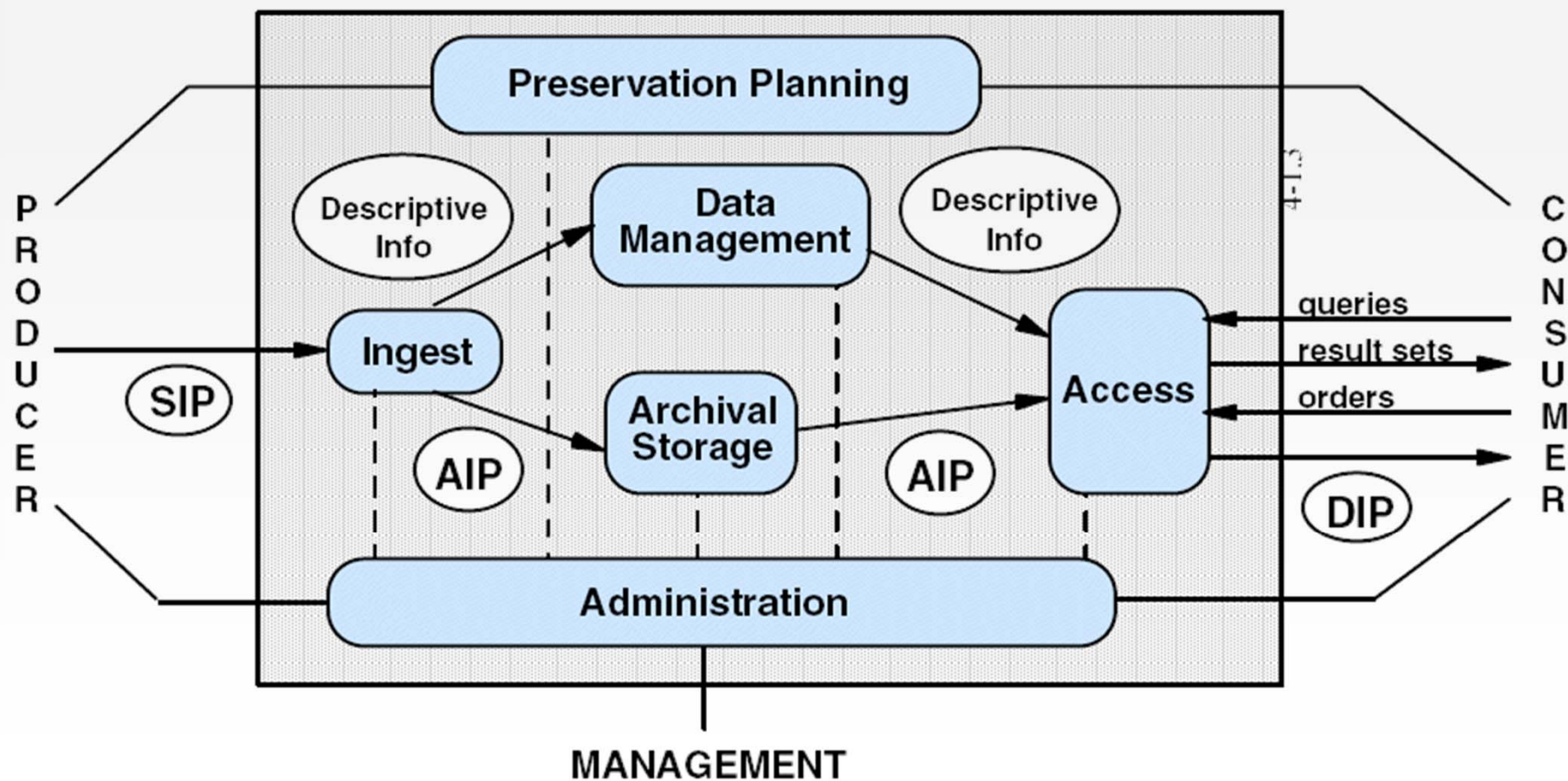
Challenge II – Data Model

Challenge III – Characterization

ExLibris
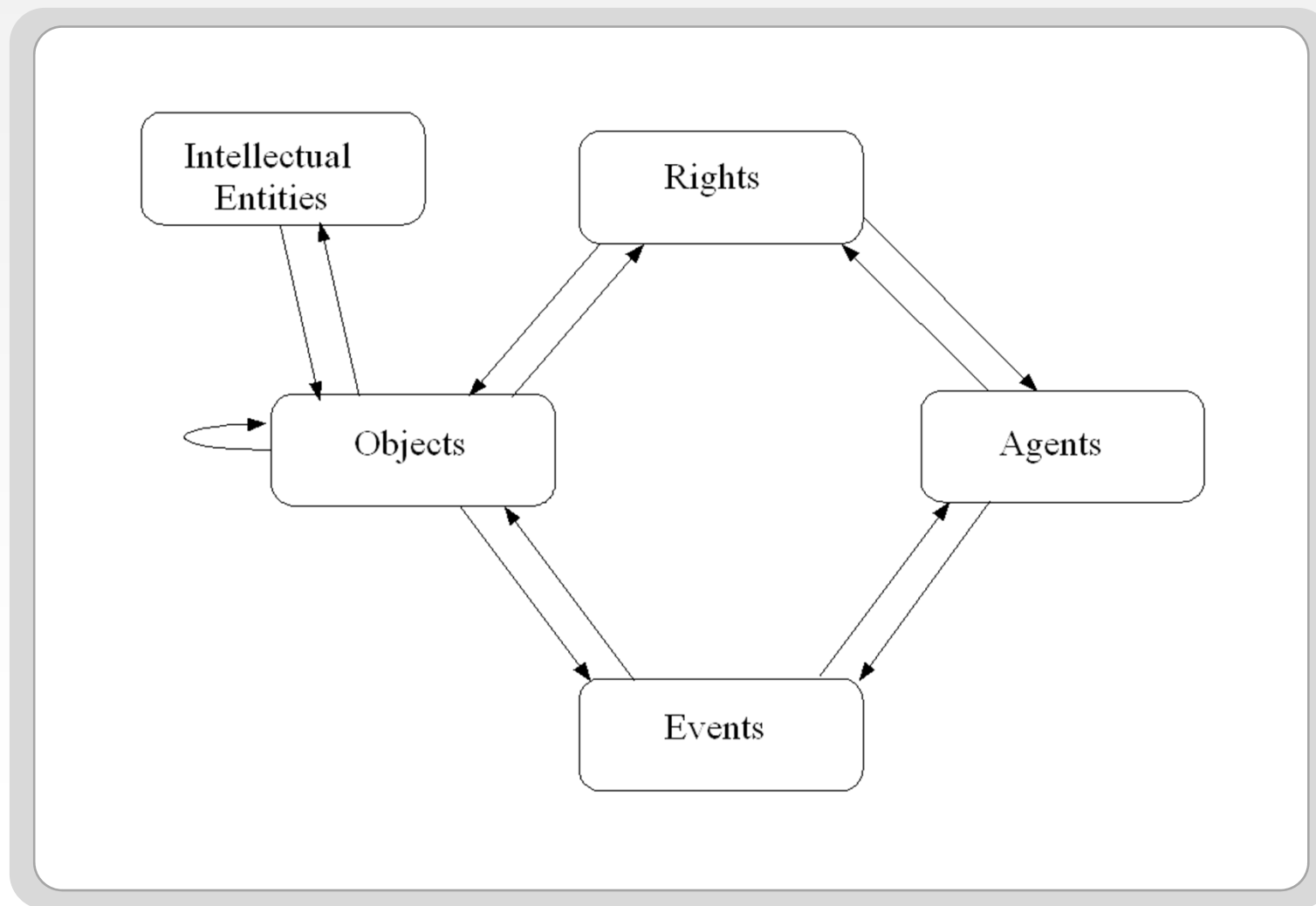
Challenge III – Characterization

- A framework around the different tools including:

  - A technical analyst workbench to handle the different issues

  - Rule based decision making to automate the process

  - Use the local format library to define rules of what to use on each format

ExLibris

Challenge IV – Existing Infrastructure

- Institutions have infrastructure in place

- Not creating yet another system

- Integrate with existing environment.

ExLibris

Challenge IV – Existing Infrastructure

Validation Plug-In **Enrichment Plug-In** **Migration Plug-In**

Rosetta

Deposit API

IE Parsing Tools

Authentication
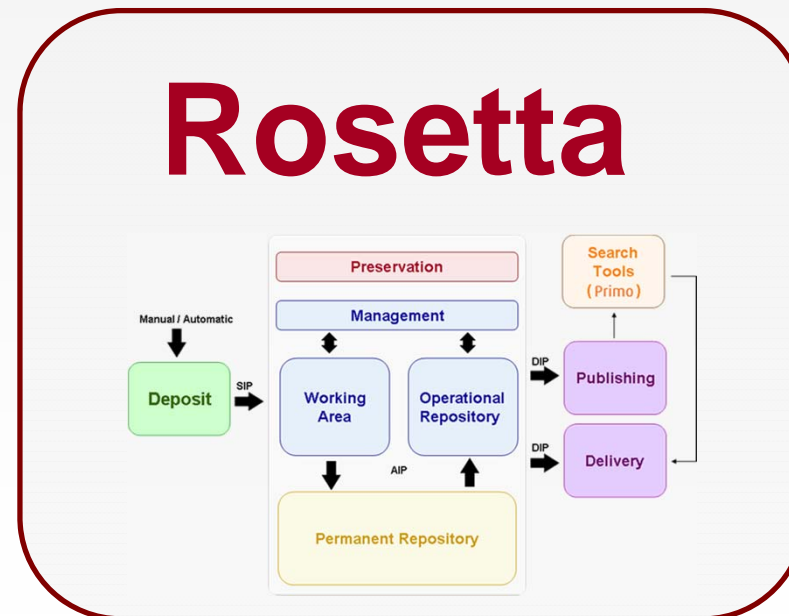
Viewers

UI Customization

SRU/SRW

OAI

Deep Links

Challenge V – The System as a Bottleneck

ExLibris

Challenge V – The System as a Bottleneck

- All preservation worthy information is stored on disk

- Storage abstraction later

- Storage is open and based on standards

- Information can be restored into Rosetta or into another system

Ex Libris
Rosetta
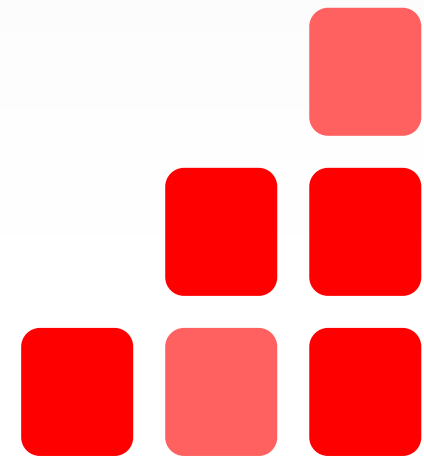
Ex Libris Rosetta

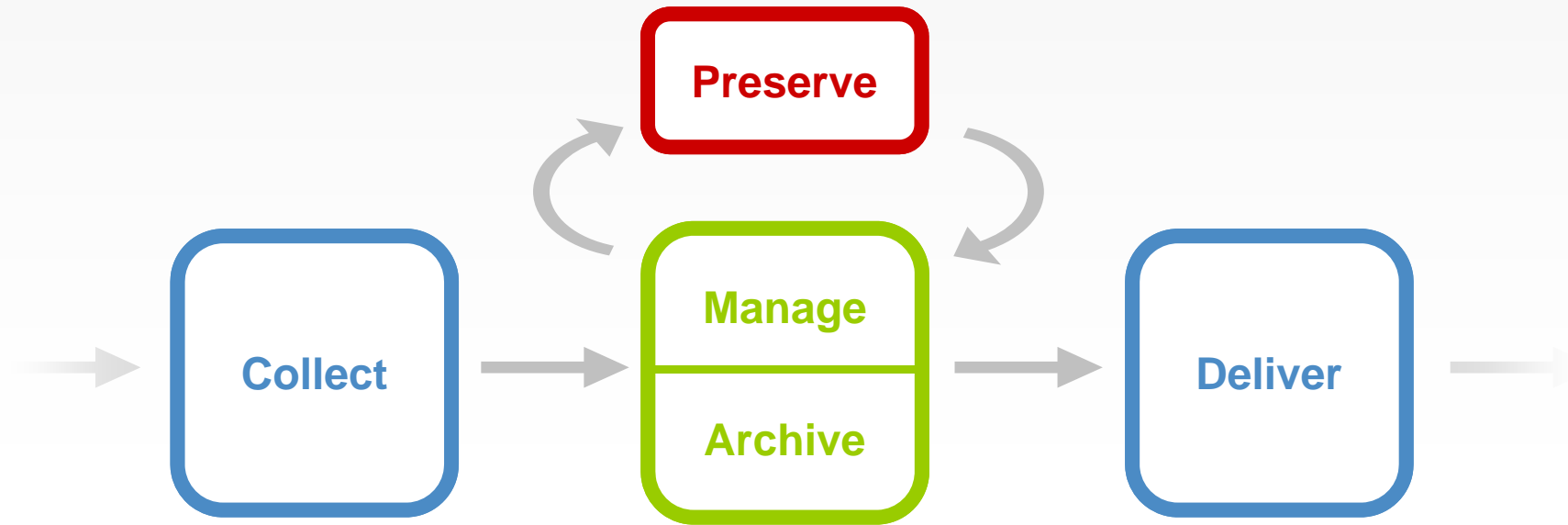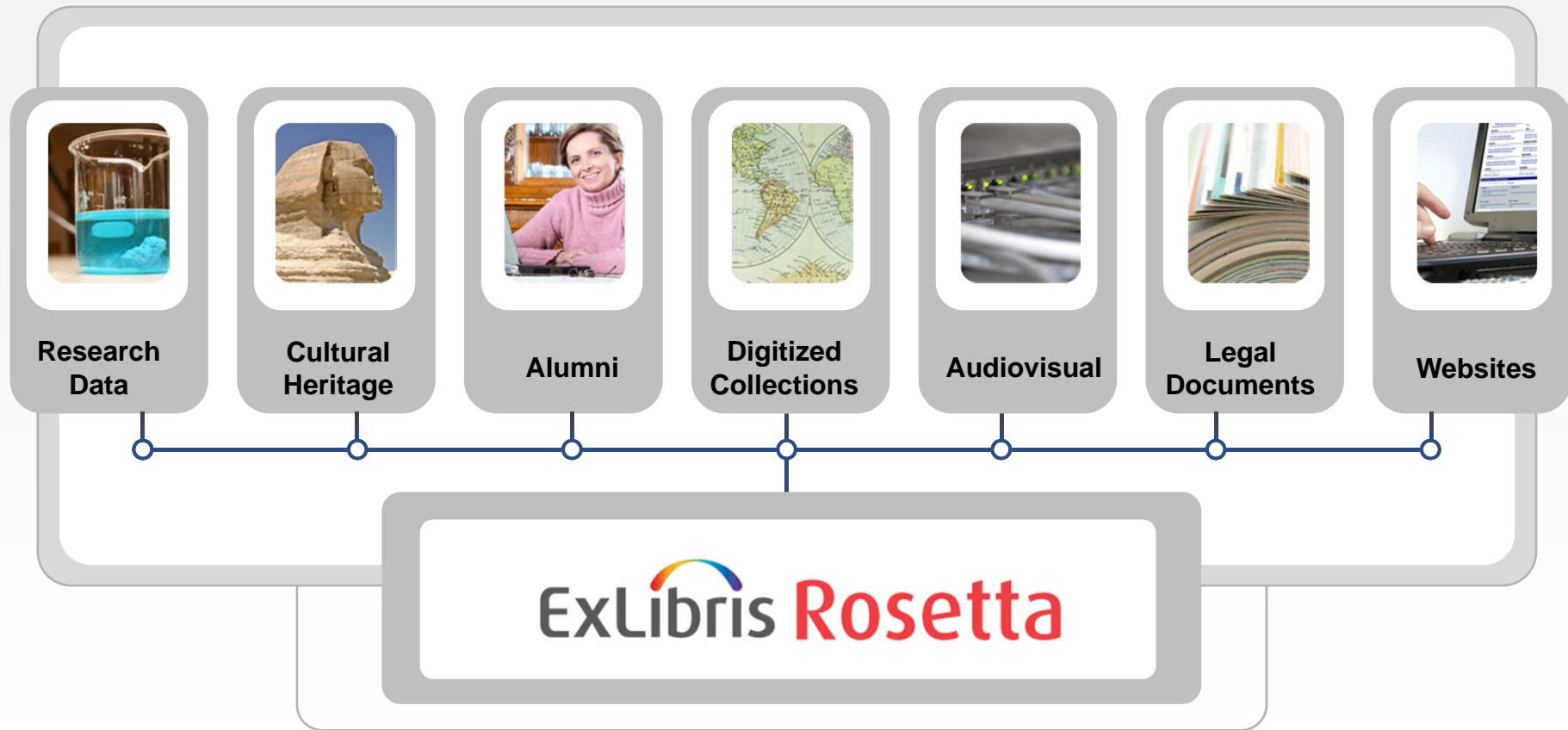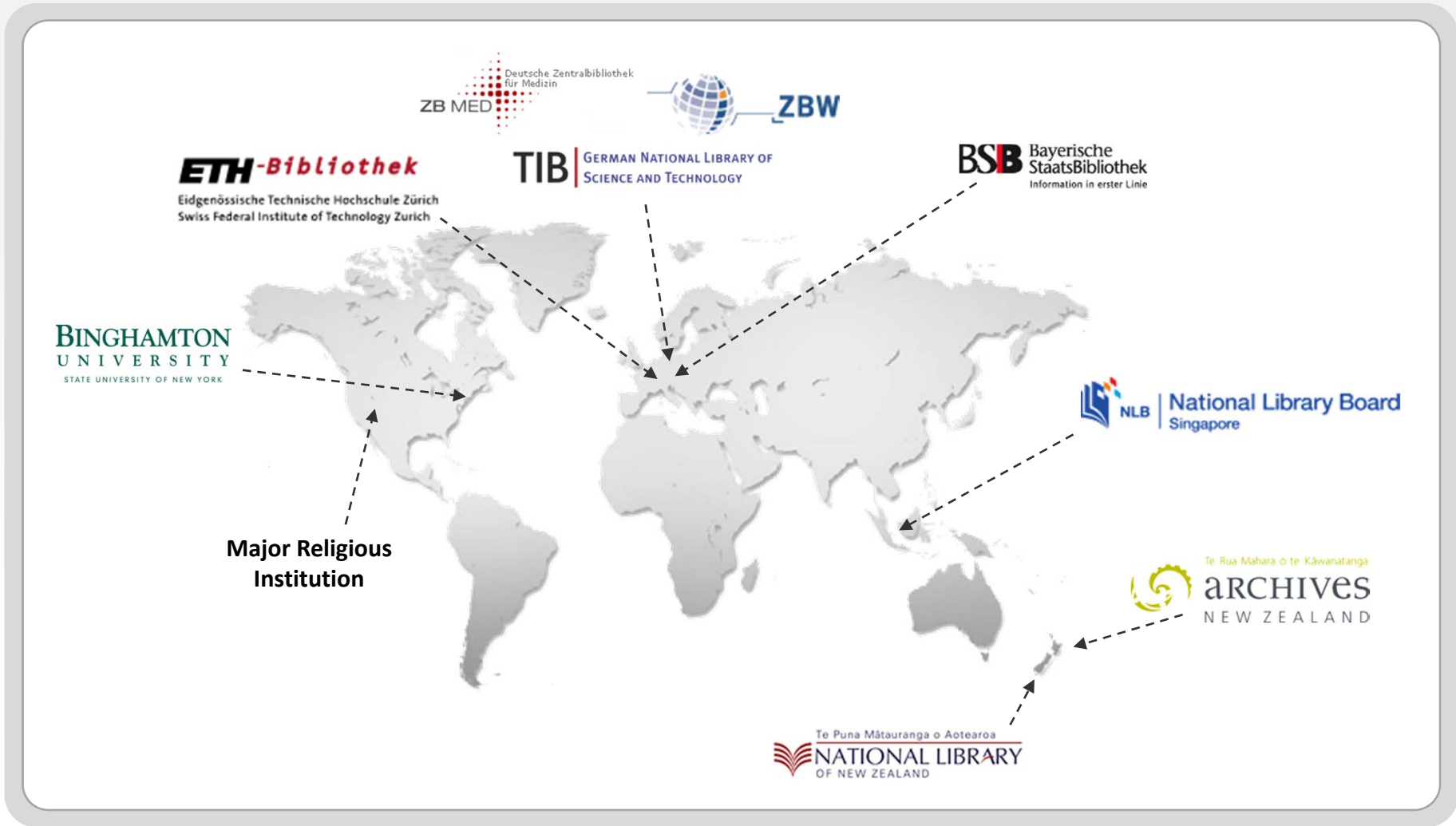**Ex Libris Rosetta** helps institutions in **collecting, managing, archiving** and **preserving** their digital collections, ensuring its **data integrity** and **access over time**

What Can Rosetta Do For You ?

Research Data | Cultural Heritage | Alumni | Digitized Collections | Audiovisual | Legal Documents | Websites

ExLibris Rosetta

ExLibris

Our Customers Around the World

Conclusions

- Digital materials are easily created and on a large scale

- Awareness for digital preservation is growing

- Challenges are well defined

<u>Rosetta</u>

- These challenges represent a small proportion of the requirements and objectives of the Rosetta digital preservation solution

- Complete preservation solution

- Scalable, flexible and extensible

Readings

Reference Model for an Open Archival Information System (2002). Available at: http://public.ccsds.org/publications/archive/650x0b1.pdf

PREMIS, Data Dictionary for Preservation Metadata(2008),

Available at: http://www.loc.gov/standards/premis/v2/premis-2-0.pdf

100 Year Archive Requirements Survey (2008),

Available at:
http://www.snia.org/forums/dmf/knowledge/100YrATF_Archive-Requirements-Survey_20070619.pdf

Sustainable Economics for a Digital Preservation, BRTF (2010),

Available at: http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf

Mind the Gap (2006),

Available at:
http://www.dpconline.org/index.php?option=com_docman&task=doc_download&gid=340
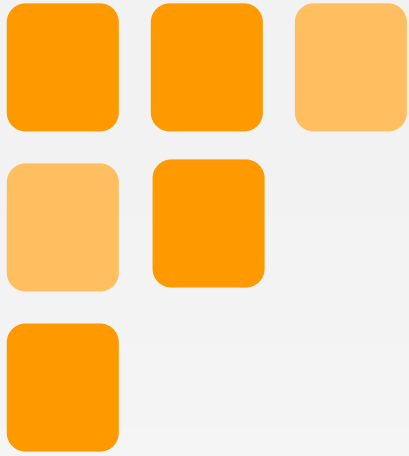
ExLibris

Thank You !!

Ido.Peled@ExLibrisGroup.com