PeDALS
Persistent Digital Archives & Library System

Richard Pearce-Moses

Deputy Director for Technology & Information Resources

Arizona State Library, Archives and Public Records

The Business Need

- Curators will have a framework
  - To describe business processes
  - That can be automated
- So that large numbers of electronic records and digital publications (if possible) can be processed
  - In a cost effective manner
  - So that the repository can work with ever-increasing quantities of materials in a time of limited resources
- One of several tools on the curator's workbench

Curatorial Rationale

- Question traditional, paper-based practices in order to transform them into the digital era
  - Appraisal
  - Acquisition
  - Arrangement and description
  - Housing and storage
  - Reference and access
  - Preservation
- Preserving archival principles of provenance, context, collective control, and authenticity and integrity

Technical Goals

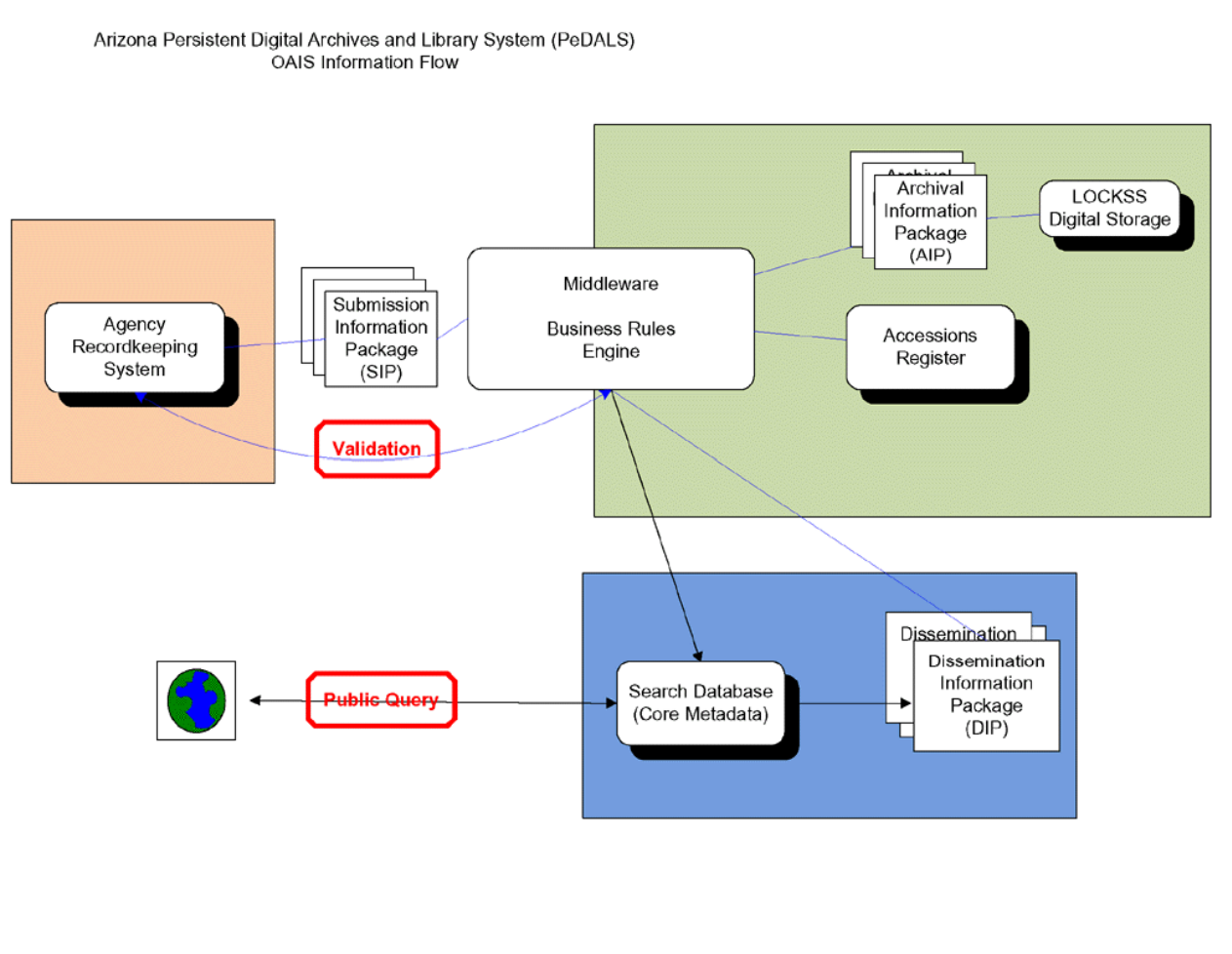
- To demonstrate the use of middleware to implement business rules in software as an integrated workflow to process collections of records and publications

- To build “digital stacks” using LOCKSS as the basis of an inexpensive storage network that can preserve the authenticity and integrity of the materials.

Additional Goals

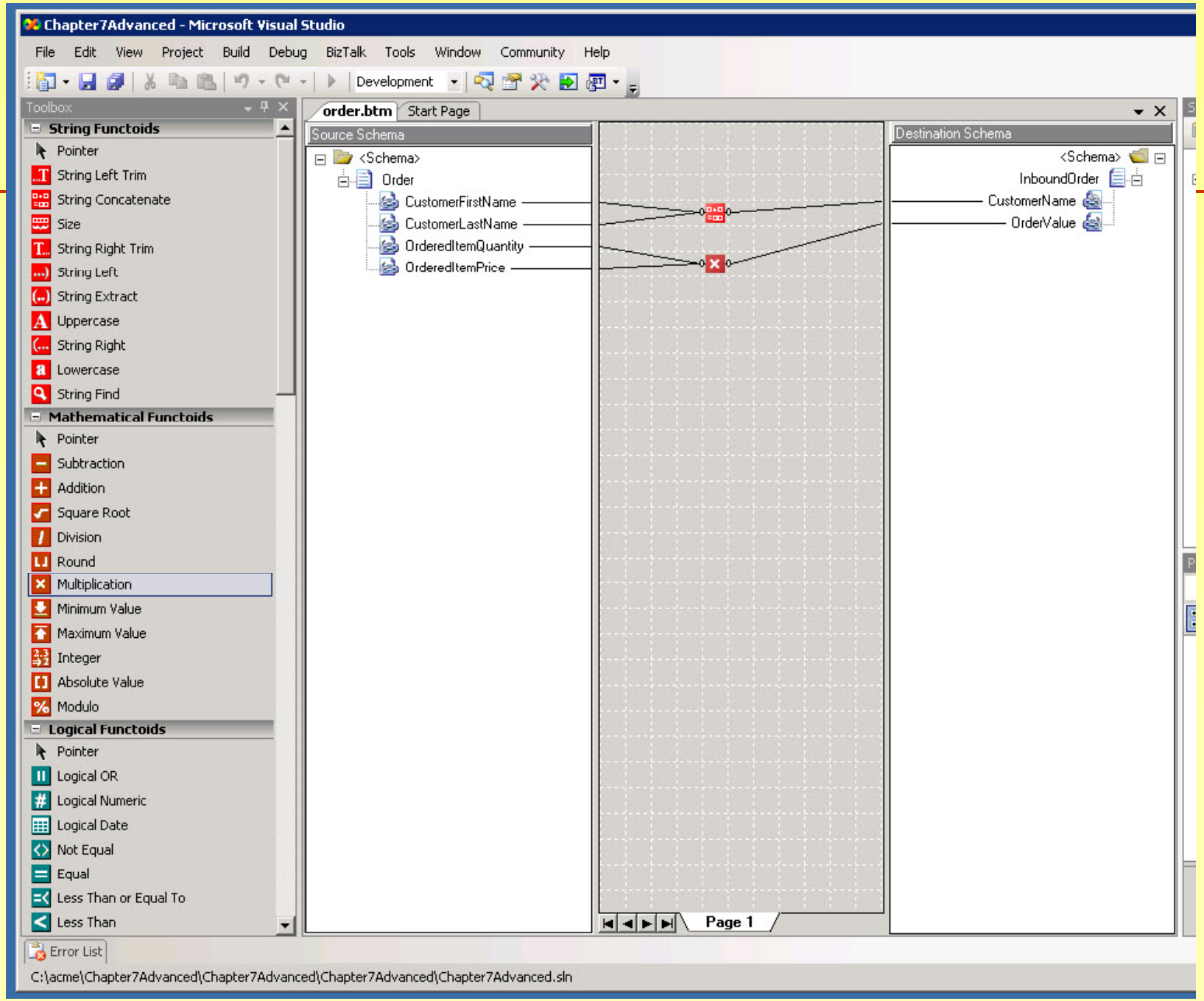
- To build a community of shared practice that meets the needs of a wide range of repositories
  - For best practices ~ what works, what's practical
  - For resource sharing ~ avoid redundant work
- To remove barriers to preservation by keeping costs as low as possible

Data Flow


Arizona Persistent Digital Archives and Library System (PeDALS)
OAIS Information Flow
Agency Recordkeeping System
Submission Information Package (SIP)
Middleware
Business Rules Engine
Archival Information Package (AIP)
LOCKSS Digital Storage
Accessions Register
Validation
Public Query
Search Database (Core Metadata)
Dissemination Information Package (DIP)

Automated Processing

- Curators work with rules, not records
  - Describe business processes (rules)
  - Monitor the process for quality assurance
- Rules expressed in software
  - A “pipeline” that transforms records as they move through the system
- Based on Microsoft BizTalk middleware


Chapter7Advanced - Microsoft Visual Studio
File Edit View Project Build Debug BizTalk Tools Window Community Help
Development
Toolbox
String Functoids
Pointer
String Left Trim
String Concatenate
Size
String Right Trim
String Left
String Extract
Uppercase
String Right
Lowercase
String Find
Mathematical Functoids
Pointer
Subtraction
Addition
Square Root
Division
Round
Multiplication
Minimum Value
Maximum Value
Integer
Absolute Value
Modulo
Logical Functoids
Pointer
Logical OR
Logical Numeric
Logical Date
Not Equal
Equal
Less Than or Equal To
Less Than
order.btm
Start Page
Source Schema
<Schema>
Order
CustomerFirstName
CustomerLastName
OrderedItemQuantity
OrderedItemPrice
Destination Schema
<Schema>
InboundOrder
CustomerName
OrderValue
Page 1
Error List
C:\acme\Chapter7Advanced\Chapter7Advanced\Chapter7Advanced\Chapter7Advanced.sln

Project Status – Completed

- Technical infrastructure installed
- Core metadata defined
- Schema for a simple AIP
- Developed administrative catalog
- AZ marriage certificates ingested, transformed and created created metadata, packaged as AIPs, and deposited in LOCKSS
- Demonstrated reuse of code by adapting rules for marriage certificates to SC Public Service Commission orders

Project Status – To Do

- Complete Administrative Catalog Interface
- Develop AIP for compound records
- Develop DIP
- Develop Public Catalog web interface
- Write rules to ingest additional records and publications
- Project to be completed by December 2009

For more information

- http://www.pedalspreservation.org/

- Principal Investigator
  - Richard Pearce-Moses
    rpm@lib.az.us

- Project Coordinator
  - Sara Muth
    smuth@lib.az.us