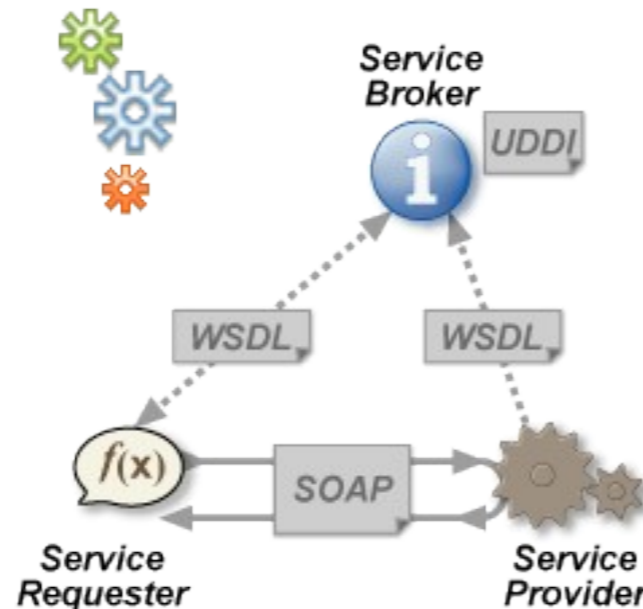


## Extending Reports into Services: New Infrastructure, Tools and Contexts

ExLibris X-Services



Kevin Kidd  
Library Applications & Systems Manager  
Boston College  
University Libraries

ELUNA 2010  
Fort Worth, TX  
May 13, 2010  
2:15 – 3:15 PM

## Agenda

- ∅ **Web Services and Library Services: Our Strategy and Requirements**
- ∅ **Administrative (e.g. Back Office) Reporting Needs**
- ∅ **Online Patron Services: (e.g. Customer-Facing) Needs**
- ∅ **Our Solution: LogiInfo and LogiAd Hoc**
- ∅ **What's Next: Future Developments**

## **Some Background...**

- ∅ **By Early 2008, the BC Libraries were in desperate need of a new Administrative Reporting System**
  - ∅ **Legacy Crystal Reports System was woefully inadequate (8 Years Old)**
- ∅ **We had recently developed a strategic plan (the Aerie Project) to build a flexible web services infrastructure to deliver new patron and back-office services**
- ∅ **The goal of the Aerie Project was to create online library services that could be combined and reused in different contexts**
  - ∅ **Library web site**
  - ∅ **Blackboard CMS**
  - ∅ **University administrative portal**
  - ∅ **Personal research portal**
  - ∅ **Personal Presence Web ([share.bc.edu](http://share.bc.edu))**

## Initially, we approached the problem on two separate tracks

- ∅ **Researched and tested multiple reporting and BI applications (Open Source and Commercial)**
  - ∅ **Jasper**
  - ∅ **Pentaho**
  - ∅ **ARC**
  - ∅ **Crystal Enterprise**
  - ∅ **LogiReports / LogiInfo**
  
- ∅ **Simultaneously began education and research on web services platforms and solutions**
  - ∅ **Began to build in-house Java programming expertise.**
  - ∅ **Mule ESB**
  - ∅ **JBoss ESB**
  - ∅ **Oracle ESB**

## Features we were looking for in a reporting solution

- ∅ **Ease of Use; Intuitive Interface**
- ∅ **Ease of Deployment (Web Browser Only; No Plugins)**
- ∅ **Scheduling of Reports**
- ∅ **Open Environment (e.g. Linux/Tomcat/Apache, etc...)**
- ∅ **Ability to integrate data from multiple sources (Aleph, SFX, MetaLib, X-Services, etc...)**
- ∅ **Reports available in multiple formats (HTML, PDF, Excel, CSV)**
- ∅ **Viability – lots of users; active developer community**
- ∅ **Reasonable license cost and terms**

## **Features we were looking for in a web services / online patron services platform**

- ∅ **Here, our initial sense was what we required was, admittedly, vague**
  
- ∅ **We wanted to be able to “glue” together unrelated data and services to build and deploy new services.**
  - ∅ **X-Services**
  
  - ∅ **Aleph, SFX, MetaLib Data**
  
  - ∅ **ILLiad and other applications**
  
- ∅ **Reusability of these new services was very important**
  
- ∅ **Open standards are very important as well (Java, Tomcat/JBoss)**

**With all of this in mind, we initially settled on**

**LogiInfo BI Platform**



## **Why?**

- ∅ **Purely web-based reports - no plugins or additional software**
- ∅ **LogiStudio: Easy to learn report development environment - no coding required**
- ∅ **Support for a Java application servers – BC uses Tomcat**
- ∅ **Access and join data from virtually any data source – SQL, Web Services, Text Files, XML, Google, Twitter, Screen Scrape, Excel, etc...**
- ∅ **Export to Excel, Word, PDF, CSV, HTML, and XML**

**But, the clincher was**

**Server-based, NOT user-based, licensing**

**Why is this crucial?**

- ∅ **Much more economical. Don't need to pay for individual users.**
- ∅ **LogiInfo potentially becomes a full-blown dynamic web application development environment, not simply a back-office reporting system.**
- ∅ **We are free to deploy applications for *everyone* at Boston College. Not just library staff.**



## **We realized that LogiInfo could actually fulfill many of the web service needs we identified in planning for the Aerie Project**

- ∅ **For example, LogiInfo's ability easily to "glue" together data from multiple, unrelated sources would allow us to build a context-sensitive Course Reserve service that:**

**Gets a Student's Current Schedule (Source: BC Scheduling System)**

**+**

**Gets the items on reserve for a given course (Source: Aleph DB)**

**=**

**New Service: My Course Reserves**

**Also, because this new My Course Reserves service was built in LogiInfo, it is highly portable - another need we identified in the Aerie plan**

- ∅ **For example, we can embed the service in Blackboard/Vista CMS, in the University Admin Portal and in course-specific LibGuides.**
- ∅ **The service can be inserted directly into an HTML page via a JavaScript “Widget”.**
- ∅ **Styles (CSS) can change on the fly to seamlessly integrate with the host web page or application.**
- ∅ **The service can mimic Aleph X-Services, and return the course reserve data as pure XML that can be used in any way the calling application requires.**

## Sample Widget – To embed in HTML. Automatically generated by LogiInfo

```
<FORM> <!-- The FORM element is required if there are INPUT elements in the Widget -->
```

```
<!-- Include this once for each page. -->
```

```
<SCRIPT src="https://arc.bc.edu:8443/LibraryAdminServices/rdTemplate/rdWidget/rdWidget.js">  
</SCRIPT>
```

```
<!-- Include this for each widget on the page. -->
```

```
<DIV ID="myWidgetDiv" >Loading...</DIV>
```

```
<SCRIPT>
```

```
var myWidget= new rdLogiWidget;  
myWidget.definition="CourseDocs";  
myWidget.containerID="myWidgetDiv";  
myWidget.setParameter("lgxPreview","41403");  
myWidget.load();
```

```
</SCRIPT>
```

```
</FORM>
```

## Sample XML for Requested In-Process Web Service

```
<rdData>
<dtRequestedInProc PATRONID="0000012826 " NAME="Ashley,Jane" EMAIL="ASHLEY@BC.EDU"
PHONE="6175524155" SYSNO="002671977" TITLE="Educating nurses : a call for radical transformation /
Patricia Benner ... [et al.] ; foreword by Le" AUTHOR="" BARCODE="B947846
"
REQ_DATE="20100506" PICKUP_LOC="ONL " REQ_ARRIVAL_DATE="20100222" calcReqDate="2010-05-
06" calcArrDate="2010-02-22" calcPickupLoc="O'Neill Library"/>

<dtRequestedInProc PATRONID="0000011418 " NAME="Bregman,Adeane" EMAIL="BREGMAN@BC.EDU"
PHONE="6175523136" SYSNO="002678481" TITLE="Der Tod - Tor zum Leben ein Trostbuch Dietrich
Steinwede" AUTHOR="" BARCODE="B924574
" REQ_DATE="20100510"
PICKUP_LOC="BAPST" REQ_ARRIVAL_DATE="20100308" calcReqDate="2010-05-10" calcArrDate="2010-
03-08" calcPickupLoc="Bapst Art Library"/>

<dtRequestedInProc PATRONID="0000011418 " NAME="Bregman,Adeane" EMAIL="BREGMAN@BC.EDU"
PHONE="6175523136" SYSNO="002683693" TITLE="The Hudson River to Niagara Falls : 19th-century
American landscape paintings from the New-York Hist" AUTHOR="" BARCODE="B927861
"
REQ_DATE="20100511" PICKUP_LOC="BAPST" REQ_ARRIVAL_DATE="20100325" calcReqDate="2010-05-
11" calcArrDate="2010-03-25" calcPickupLoc="Bapst Art Library"/>

<dtRequestedInProc PATRONID="0000012826 " NAME="Ashley,Jane" EMAIL="ASHLEY@BC.EDU"
PHONE="6175524155" SYSNO="002671977" TITLE="Educating nurses : a call for radical transformation /
Patricia Benner ... [et al.] ; foreword by Le" AUTHOR="" BARCODE="B947846
"
REQ_DATE="20100506" PICKUP_LOC="ONL " REQ_ARRIVAL_DATE="20100510" calcReqDate="2010-05-
06" calcArrDate="2010-05-10" calcPickupLoc="O'Neill Library"/>
</rdData>
```

# Requested In-Process Barcode Flags Generated from XML Service

The screenshot shows a Mozilla Firefox browser window with the URL <http://arc.bc.edu:8080/PrintFlags/PrintFlags.do>. The page displays four columns of 'RUSH IN PROCESS' barcode flags. Each flag contains the following information:

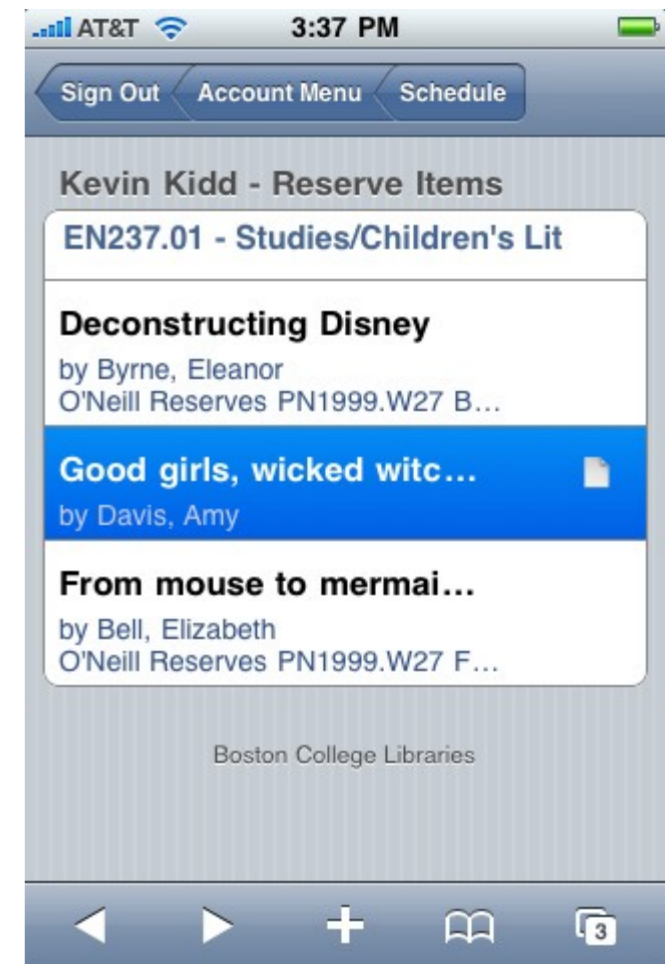
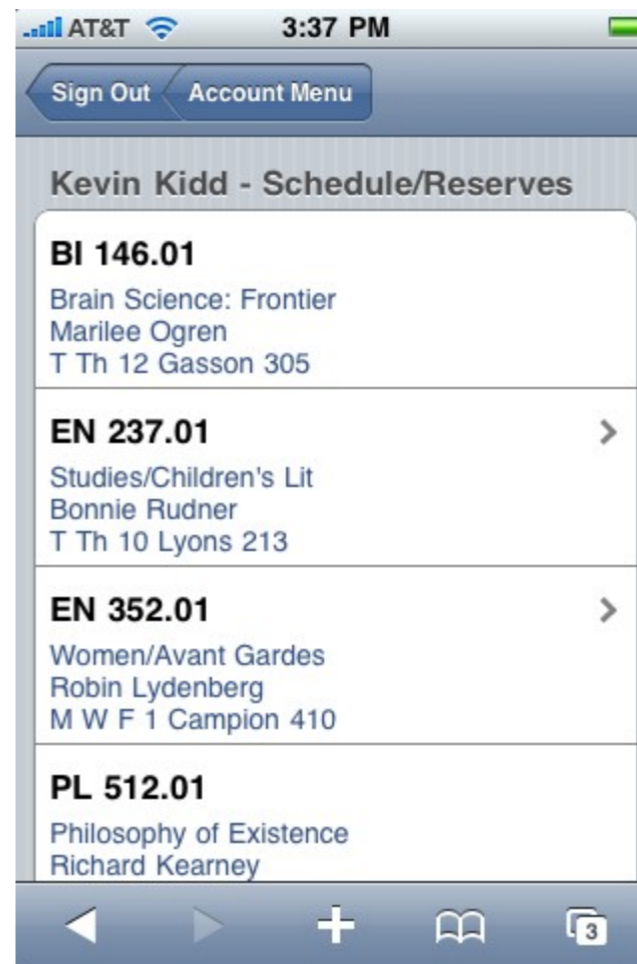
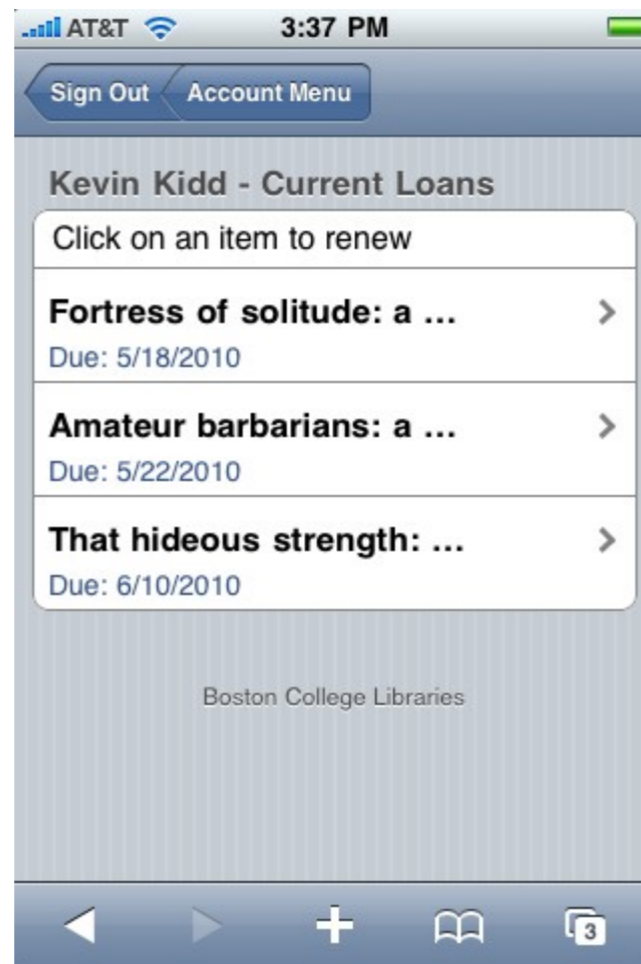
- Header:** RUSH IN PROCESS
- Arrival Date:** (e.g., 2010-02-22, 2010-03-08, 2010-03-25, 2010-05-10)
- Barcode:** (e.g., B947846, B924574, B927861, B947846)
- Author:** (e.g., Patricia Benner, Dietrich Steinwede, 19th-century American landscape paintings from the New-York Hist)
- Title:** (e.g., Educating nurses : a call for radical transformation / Patricia Benner ... [et al.] ; foreword by Le; Der Tod - Tor zum Leben ein Trostbuch; The Hudson River to Niagara Falls : 19th-century American landscape paintings from the New-York Hist)
- Patron / Requestor Info:** (e.g., Aleph Patron ID: 0000012826, Name: Ashley, Jane, E-Mail: ASHLEY@BC.EDU, Phone: 6175524155, Request Date: 2010-05-06, Pickup Location: O'Neill Library)
- System No:** (e.g., 002671977, 002678481, 002683693, 002671977)
- Comments:** (empty text area)

The browser's taskbar at the bottom shows several open applications, including 'onei212kiddk2.bc.e...', 'C:\Users\kiddk\Des...', 'Microsoft PowerPoi...', 'test\_logi\_widget.ht...', and 'Print Flags - Mozilla ...'. The system clock indicates the time is 6:22 PM.

## **What else can we do with LogiInfo?**

- ∅ **“My Library”-like portal (using Dashboards)**
- ∅ **Library Resource Recommender (based upon Major, a patron’s course schedule, or other user-selected criterion)**
- ∅ **Easily build interfaces to Aleph, MetaLib and Primo X-Services and/or SOAP Services, including view/renew items, search the catalog, etc**
- ∅ **Build a full-blown Mobile Library Web site (why not?)**

## Mobile Library: Built in LogiInfo



**BUT – we quickly understood we needed some major infrastructure to fully realize this vision**

- ∅ The Aleph application database is not built for reporting (not even close)
- ∅ It is even less suited to support linking to data in other databases
- ∅ The potential that hundreds of additional users (e.g. patrons) might be accessing Aleph data on a daily basis (though LogiInfo) had worrisome performance implications.
- ∅ The upshot – We needed a data warehouse for reporting and application purposes.
- ∅ ARC would not work because we wanted to integrate data from sources such as SFX, MetaLib and Verde

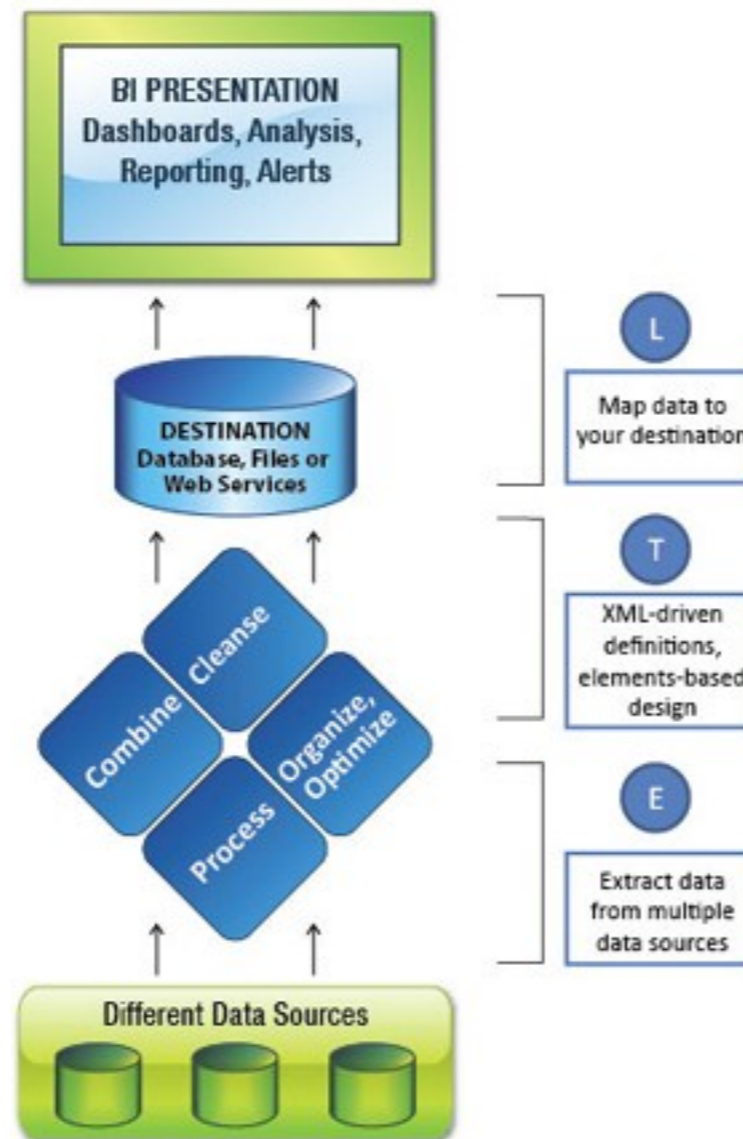


## The Solution: LogiETL, LogiAdHoc and the Data Warehouse

- ∅ In mid-2009, we started discussions with LogiXML.
- ∅ In exchange for LogiETL and LogiAdHoc licenses, BC agreed to help LogiXML developers build a generalized data warehouse.
- ∅ Initially, we are focusing on a warehouse for the Aleph database
- ∅ The warehouse can be extended to include data from SFX, MetaLib, or, indeed any other related data source

## The Aleph Data Warehouse Software and Architecture

- ∅ LogiETL: Used to extract, transform/normalize and load Aleph Data into the data warehouse



# LogiETL Development Studio

The screenshot displays the LogiXML ETL - Logi ETL Studio interface. The main workspace is titled 'ItemDim' and contains a hierarchical tree of ETL components. The tree structure is as follows:

- ItemDim
  - Dimensions.ItemDim
    - SetUpDateLookupValue
    - Session Parameters
      - Set this to -1 to perform an initial load.
      - latest Z30\_UPDATE\_DATE is 20100201
    - Id\_data\_save\_file\_name
      - dl\_static\_data\_save\_file\_name
        - Static Data Row
    - prcifPerformInitialLoad
      - Issues
      - Record Count: 2,507,803
      - Total Job Execution Time: ~ 14,060 s
      - extItemDim
        - dIBCL50.Z30
          - Query Run Time (Toad): ~
          - Time for database server to return data: ~ 11,721 s
          - Datalayer Creation Time: ~ 11,911 s
        - ddupITEM\_NUMBER
          - Items Removed: 0
        - GOV\_DOC\_BARCODE
          - Temporary calculated column to work around Logi ETL bug with update of Boolean fields.
          - This is not really needed with the bulk load, but is here for consistency.
      - Save\_To\_CSV
      - ise\_CSV\_Copy\_to\_DB
      - sql\_bulk\_load\_data
        - Execution Time: ~ 95 s
        - Records Loaded: 2,507,803
      - prcifPerformIncrementalUpdate
      - extNewItemDimRecords
        - dIBCL50.Z30
          - ddupITEM\_NUMBER
          - GOV\_DOC\_BARCODE
            - Temporary calculated column to work around Logi ETL bug with update of Boolean fields.
        - ldtUpsert\_ItemDim
          - Execution Time (batch 1000 -- 1,000 records): ~ 5 s
          - Execution Time (batch 1000 -- 10,000 records): ~ 35 s
          - Execution Time (batch 1000 -- 100,000 records): ~ 350 s
          - cmItem\_Number
          - cmCall\_Number
          - cmCall\_Number\_Type
          - cmBarcode
          - cmCollection
          - cmMaterial
          - cmLC\_Class
          - cmClass Letters

The right-hand pane is titled 'Attributes - EtlJob' and contains an 'Attribute Spy' section with a table of attributes:

*Required	
ID	Dimensions.ItemDim

Below this, there is an 'Optional' section with a text field for 'ETL Job Description'. The 'ETL Job Description' section contains the text: 'Double click an Attribute Name for Attribute Zoom.' Below that is an 'Element Toolbox' with various ETL components like 'Email.Log', 'Execute.Sql', 'FTP.Upload', etc.

## The Aleph Data Warehouse Software and Architecture

- ∅ **Data Warehouse: PostgreSQL (Open Source Database)**
- ∅ **AdHoc Server: VMWare Virtual Linux Server (easy to bundle and distribute)**
- ∅ **LogiAdHoc: Allows non-techies to easily build reports as they require.**
  - ∅ **No client installation. Interface is purely web-based**
  - ∅ **Users can create charts, graphs and dashboards**
  - ∅ **Reports can be scheduled and automatically delivered**

**DEMO**

# Thank You

- Questions?
- contact: [kevin.kidd@bc.edu](mailto:kevin.kidd@bc.edu)