Integrated Decision Support System – iDSS for Library Holistic Evaluation

Lorena Siguenza-Guzman
Victor Saquicela
Dirk Cattrysse

3 June 2014
Dynamic Information Environment

- e-content and data growth
- New emerging technologies
- Diverse user needs

Knowledge Management a powerful tool for libraries

High-quality Services at Lower Costs

# Holistic Framework Matrix

<table>
<thead>
<tr>
<th>Perspective</th>
<th>Topic</th>
<th>Collection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Internal (Library)</strong></td>
<td>Library System</td>
<td>How is the library system manipulated?</td>
</tr>
<tr>
<td></td>
<td>What does the library system consist of?</td>
<td></td>
</tr>
<tr>
<td><strong>External (User)</strong></td>
<td>How effective is the library system?</td>
<td>How useful is the library system?</td>
</tr>
</tbody>
</table>

Holistic Data Collection

**Library System**

- **Cost Analysis Processes**
  - Processes, Time, Resources

- **Quality**
  - Statistics gathering, Suggestion boxes, Usability testing, Satisfaction surveys

**Collection**

- **Log analysis**
  - Implicit and explicit data

- **Bibliometrics**
  - Citation patterns, Publishing patterns, Journals downloaded, and Journal’s impact factor

---

First Quadrant - Costing Analysis

**Library System**
- Cost Analysis Processes
  - Processes, Time, Resources

**Collection**
- Log analysis
  - Implicit and explicit data

**Internal Perspective (Library)**
- Cost Analysis Processes

**External Perspective (Users)**
- Quality
  - Statistics gathering, Suggestion boxes, Usability testing, Satisfaction surveys
- Bibliometrics
  - Citation patterns, Publishing patterns, Journals downloaded, and Journal’s impact factor
First Quadrant - Costing Analysis

• Time-Driven Activity-Based Costing

First Quadrant - Costing Analysis

**Acquisition**
- Acquisition of Books
- Acquisition of Journals

**Cataloging**
- Original Cataloging
- Copy Cataloging
- Cataloging Closed-Stack Items

**Circulation**
- Lending items
- Returning items
- Reference

**Document Delivery**
- Requesting closed stack items
- ILL Outgoing Request
- ILL Incoming Request – digital items
- ILL Incoming Request – printed items
Second Quadrant - Quality

Internal Perspective (Library)
- Cost Analysis Processes
  - Processes, Time, Resources

External Perspective (Users)
- Quality
  - Statistics gathering, Suggestion boxes, Usability testing, Satisfaction surveys

Collection
- Log analysis
  - Implicit and explicit data

Library System
- Bibliometrics
  - Citation patterns, Publishing patterns, Journals downloaded, and Journal’s impact factor
Second Quadrant - Quality

- Assessment tools
  - Statistics gathering
  - Suggestion boxes
  - Web usability testing
  - User interface usability
  - **Satisfaction surveys**
  - ...

![LibQUAL+®](image-url)
Third Quadrant - Usefulness

**Internal Perspective (Library)**
- **Library System**
  - Cost Analysis Processes
    - Processes, Time, Resources

**External Perspective (Users)**
- **Collection**
  - Log analysis
    - Implicit and explicit data

- **Quality**
  - Statistics gathering, Suggestion boxes, Usability testing, Satisfaction surveys

- **Bibliometrics**
  - Citation patterns, Publishing patterns, Journals downloaded, and Journal’s impact factor
Third Quadrant - Usefulness

- Citations patterns
- Publishing patterns

Analysis

EBSCO

Scopus
Fourth Quadrant - Users’ behavior

- **Library System**
  - **Cost Analysis Processes**
    - Processes, Time, Resources
  - **Quality**
    - Statistics gathering, Suggestion boxes, Usability testing, Satisfaction surveys

- **Collection**
  - **Log analysis**
    - Implicit and explicit data
  - **Bibliometrics**
    - Citation patterns, Publishing patterns, Journals downloaded, and Journal’s impact factor

**Internal Perspective (Library)**

**External Perspective (Users)**
Fourth Quadrant - Users’ behavior

Bibliographic Surrogate
- MARC, Dublin Core
- Author, Publisher
- Subject headings/
  Classification
- Abstract, Keywords
- Type of material
- Location (Physical, Virtual)
- Accessibility

Bibliometric Data
- Requires non-aggregated e-resource data
- Citations / Links
- Social networks, e.g.
  Disciplines
  Affiliations

Library Services
- Basics: Time, Date, Location, Format, Concurrent use

<table>
<thead>
<tr>
<th>Searching</th>
<th>Circulation / Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content (flag for failed search)</td>
<td>OPAC info (circ lengths)</td>
</tr>
<tr>
<td>Path of search (before/after)</td>
<td>Transaction Log Analysis</td>
</tr>
<tr>
<td>Resolution of search</td>
<td>Click through link / Proxy Info</td>
</tr>
<tr>
<td>Concurrent use of materials</td>
<td>COUNTER data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>Outreach and Training</th>
<th>ILL / Other Requests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>Program type</td>
<td>Stream of request</td>
</tr>
<tr>
<td>Question</td>
<td>Program title</td>
<td>Final supplier</td>
</tr>
<tr>
<td>Question topic</td>
<td>Program subject</td>
<td>Time of processing</td>
</tr>
<tr>
<td>Time taken</td>
<td>Attendee demographics</td>
<td>Time to supply</td>
</tr>
<tr>
<td>Answer</td>
<td>Requestor information</td>
<td>Costs</td>
</tr>
<tr>
<td>Referral Path</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Demographic Surrogate
- Depends upon setting:
  College / Department / Level
  Zip code / Group membership /
  Areas of Interest
- Department / Company Role
  Building / Grade / Teacher
A holistic approach to supporting academic libraries in resource allocation processes. The Library Quarterly: Information, Community, Policy.
DSS architecture – UC library

Data source
- ABCD
  - ISIS, MySQL
- TDABCD
  - MySQL
- DSPACE
  - PostgreSQL
- HR
  - Oracle/DB2
- Inventory
  - Oracle
- Academic
  - Oracle
- Survey
  - MySQL
- EZ Proxy

ETL process
- Data Integration Pentaho
  - Kettle
- Cubes Schema Workbench
- Data Warehouse
  - MySQL

Data presentation
- Online Analytical Processing
  - Saiku Pentaho
- Biblimining
  - WEKA
- Web Browser
  - Internet Explorer, Firefox, Chrome, ...

Data Flow

Requirement analysis
Data source analysis
Logical model
Data integration
Requirement Analysis

• Examples of requirements:

1. What is the number of loans of a particular item, of a given author and title, of a specific item category of a particular librarian in a time unit?

2. What is the number of loans of a particular librarian of a specific library branch in a time unit?

3. What are the operating costs of lending services of a particular library branch in a time unit?

4. What is the number of returns of a particular librarian of a specific library branch at a given user type in a time unit?

5. What is the number of fines and its corresponding value in a time unit?
Requirement Analysis

• Indicators and perspectives of the lending process

1. What is the number of loans of a particular item, of a given author and title, of a specific item category of a particular librarian in a time unit?
# Data Source Analysis

## UC library data sources

<table>
<thead>
<tr>
<th>Application</th>
<th>Modules</th>
<th>Database</th>
<th>Type of database</th>
<th>Main attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Internal data sources</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCD</td>
<td>Cataloguing</td>
<td>ISIS</td>
<td>Documental</td>
<td>Item ID, Author(s), Title, …, Cataloguer</td>
</tr>
<tr>
<td></td>
<td>Circulation</td>
<td>MARC21</td>
<td></td>
<td>Lending, Returning, Fines, Reserves, Interlibrary loan</td>
</tr>
<tr>
<td></td>
<td>TD-ABC-D</td>
<td>MySQL</td>
<td>Relational</td>
<td>Activities, Resources, Responsible, Processes, Times, Costs</td>
</tr>
<tr>
<td></td>
<td>Reference Satisfaction Survey</td>
<td>MySQL</td>
<td>Relational</td>
<td>Librarian, Inquiry, Solution, Satisfaction level, PC Terminal, Survey</td>
</tr>
<tr>
<td>LibQUAL+®</td>
<td>Quality survey</td>
<td>MySQL</td>
<td>Relational</td>
<td>Campus, Questions, User, Time</td>
</tr>
<tr>
<td>DSpace</td>
<td>Institutional Repository</td>
<td>PostgreSQL</td>
<td>Relational</td>
<td>Community, Collection, Item, Bundle, Bitstream, Bitstream Format</td>
</tr>
<tr>
<td></td>
<td>EZproxy</td>
<td>Remote Access</td>
<td>LogFile</td>
<td>IP address host accessing, Date/time of request, URL Requested, Method of request, # bytes transferred</td>
</tr>
</tbody>
</table>

## University data sources

<table>
<thead>
<tr>
<th>Application</th>
<th>Modules</th>
<th>Database</th>
<th>Type of database</th>
<th>Main attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olympo</td>
<td>Inventory</td>
<td>Oracle</td>
<td>Relational</td>
<td>Fixed assets, Class, Section, Area, Address, Provider, Fixed asset document</td>
</tr>
<tr>
<td>GSocioeconomic</td>
<td>Socioeconomic Data</td>
<td>Oracle</td>
<td>Relational</td>
<td>Entry/Exit students, Family members, Academic period, Student information, Faculties, Academic status</td>
</tr>
<tr>
<td>Academic</td>
<td>Enrollment</td>
<td>Oracle</td>
<td>Relational</td>
<td>Enrollment, Career, Academic period, Subject</td>
</tr>
<tr>
<td>HRM</td>
<td>Human Resources</td>
<td>Oracle</td>
<td>Relational</td>
<td>Personal information, User type, Department</td>
</tr>
</tbody>
</table>

## External data sources

<table>
<thead>
<tr>
<th>Application</th>
<th>Modules</th>
<th>Database</th>
<th>Type of database</th>
<th>Main attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scopus reports</td>
<td>Citation database Vendor-supplied statistics</td>
<td>Excel</td>
<td>Spreadsheet</td>
<td>Title, Publisher, 5-year Impact Factor, Topic</td>
</tr>
<tr>
<td>EBSCO statistics</td>
<td>Excel</td>
<td>Spreadsheet</td>
<td>Year, Month, Searches, Total full text, PDF full text, HTML full text, Image/Video, Abstract</td>
<td></td>
</tr>
</tbody>
</table>
Data Source Analysis

• UC library data sources
Data Source Analysis

• Conceptual data model
Logic Model

• Logic model of the lending process

1. What is the number of loans of a particular item, of a given author and title, of a specific item category of a particular librarian in a time unit?
Logic Model

- Multidimensional model of the lending process
Data Integration

• Extracting data

• Cleansing and transforming data
  – First surname, Names
  – First surname Second surname, Names
  – First surname Second surname, First name
  – First surname, First name initial

• Loading data
Data presentation

- Data reporting, OLAP and bibliomining tools

1. What is the number of loans of a particular item, of a given author and title, of a specific item category of a particular librarian in a time unit?
What’s next???

• Use of Bibliomining techniques

• Optimization models to analyze specific scenarios

• Incorporation of additional data sources
  – the syllabus management system
  – citation analysis, benchmarking
  – Web portal statistics

• Use of semantic technologies
Thanks for your attention

Lorena Siguenza-Guzman
Research Assistant
Centre for Industrial Management
Faculty of Engineering - KU Leuven
Lorena.SiguenzaGuzman@kuleuven.be