

# Institutional Archiving

Vancouver Island University  
Nanaimo, BC Canada

Andrew Speed  
Manager, Systems Development  
[andrew.speed@viu.ca](mailto:andrew.speed@viu.ca)

# What is Institutional Archiving?

- The process of preserving the history of the institution.
  - Traditionally done by placing boxes of files into storage.
  - Recalling information a manual process, tedious and time consuming.
  - Risk significant loss of history from fire/flood/misplacement of files.
  - Overwhelming when your history dates back over many decades.
- VIU archiving goals...
  - Digital storage of all paper documents.
  - Redundant offsite backups for disaster recovery.
  - Ability to search all or portions of archive rapidly.
  - Beyond archiving, a warehouse or store of current institutional documents.

# Imaging history @ VIU

- 2000 - Developed in-house scanning system/imaging archive/web viewer for student records.
- 2002 – Developed scanning system/imaging archive/web viewer for financial system.
- 2004 – Developed text searchable Archive for Board of Governors. A combination of paper documents from 1968 – 1995, and electronic documents from after 1995. Scanning process included OCR (optical character recognition) to make documents keyword searchable along with 3<sup>rd</sup> party indexing software.
- 2005 Many other committees added.
- 2008 – Scanning system/imaging archive/viewer for HR. First phase having all employee files archived.

# Issues....

- Now that archiving is fairly common at VIU, many departments and committees want on the bandwagon. Can't keep up to demand for new archives.
- Cumbersome maintenance for many similar, yet slightly different archives.
- Even though current archives are secure through existing network security, it is still possible for IT staff with privileged accounts to browse restricted archives.
- Tedious to maintain the security groups for all the different archives.

# The dream....

- Develop an institutional archiving system that can be leveraged by any department or committee without any custom programming or database setup.
- Departments can setup and maintain archive security themselves.
- Archive is text searchable without the need for any 3<sup>rd</sup> party indexing software.
- Archive is highly secure, and not accessible on the network file servers.
- Users can define what types of documents they want to store in the archive.
- Archive should handle most common file types. Ie: PDF, DOC, DOCX, TXT, RTF, HTM, XLS, JPG, GIF, PS, XML, MSG, TIF, BMP
- All UI components of the system are web based.

# The solution....

- Store files directly in database in BLOB (Binary Large Object) fields, rather than on network server.
- Leverage existing functionality in Oracle that allows indexing of BLOB's (for text searchable capability).
- Write a web based application that allows users to create archives, manage security, upload and maintain the archive, and define their own document types.

# The benefits....

- Enhanced security. Privileged IT users can no longer “browse” archives on network file server.
- Cleaner application. All archive information now in the same place, not spread across database and network file servers.
- Users can load, maintain and search their archives from any computer with an internet connection.
- In some cases could replace department “shared drives” for common file storage. No need to maintain all these network security groups for file sharing.
- Improved performance. Old system required image to be copied from secure file server to a temporary viewing area, then opened in the browser. The new system streams the document directly to the browser, eliminating the need for a file copy on the server.
- More leisure time for myself!

# Meeting vs Document Archive

- Project was initially designed to house committee meeting documents ie: Agendas, Minutes, Reports, Handouts and Information Items.
- It soon became apparent that users wanted to archive non-meeting related documents.
- An "Archive type" parameter was added to differentiate between meeting and document archives. The main difference being how the archive is sorted in the drilldown tree.
- Meetings are organized by Date, document archives are organized by document type.



# Conceptual hierarchy.....

## Archive Group

Area or departmental grouping of archives.

## Archive

Security Model: Public/Employee only/Restricted

Archive Type: Meeting Archive or Document Archive

Administrators: Users who can add documents, maintain security and document types within the archive.

Users: Can view and search in this archive.

## Document

Description: Description of document.

Document Type: User definable type of document.

Document Date: Date associated with document.

# Web Interfaces

- Archive Maintenance Tool. Allows archive “administrators” to...
  - Create new archives within their archive group.
  - Maintain archive security, adding and removing users.
  - Create or add new document types.
  - Add, change or remove documents from an archive.
- Restricted Viewer
  - Allows users with viewing access to restricted or employee access archives to view documents via drilldown or keyword search.
  - Requires users to login with network account
- Public Viewer
  - Allow general public access to open archives via drilldown or keyword search.

# Development components

- Oracle 10g Standard Edition License
- Microsoft IIS 6.0 web server
- Dhtmlx javascript library – freeware version
- Coded in ASP and javascript
- Total development time – 2 weeks