Using Microsoft Pivot to Show Genealogical Content

Cameron Engh and Jake Gehring
enghch@familysearch.org | gehringjg@familysearch.org

Overview

Pivot is a dynamic faceted search that helps the user visualize a collection of items and quickly browse and analyze it by selecting "facets." Pivot's strengths are the speedy display of hundreds of images, with a simple user interface to support the broadening or narrowing of search criteria or collection filtering.

All the tools, documentation , SDKs and APIs needed to create and display pivot collections are free downloads from Microsoft. (Overview information and downloads can be found at http://www.silverlight.net/learn/pivotviewer/)

Structure and Transport

Pivot data must be hosted on a web server. Since the data are stored in flat files, Pivot does not require any specific web server. Microsoft IIS, Apache, or any other HTTP web server support Pivot. For more information on hosting see http://www.silverlight.net/learn/pivotviewer/collection-hosting/.

Pivot's storage structure is relatively simple. All the metadata for a Pivot collection is contained in a single XML file (usually with a file extension of .cxml). The schema of this file is a published standard. More information can be found at http://www.silverlight.net/learn/pivotviewer/collection-xml-schema/).

Pivot relies on DeepZoom (Microsoft's technology formerly known as SeaDragon: http://zoom.it/) to provide a smooth zooming experience. For the Silverlight PivotView experience this is done with the built-in MultiScaleImage control. Because of this, the images for a Pivot collection are stored in the Deep Zoom hierarchical tile file structure. This consists of a root-level XML file (usually with a .dzc extension) that refers to a set of nested image-level XML files (usually with a .dzi extension). Each of these files keeps track of the many image tiles that make up an image pyramid for a single item in the collection. See http://www.silverlight.net/learn/pivotviewer/collection-image-content/ for more details. Pivot uses HTTP as its transport layer. This makes it available through the web, not blocked by firewalls or other online hurdles.

There are two options to view a Pivot collection, the Silverlight Pivot Viewer and an installable desktop product. (The desktop product, however, may be discontinued in the future.) The Silverlight PivotViewer Control is a free download, as mentioned above: http://www.microsoft.com/downloads/en/details.aspx?displaylang=en&FamilyID=9a1bb862-d80c-4145-9320-b279a63bff91).

Designing a Collection in Pivot

There are three flavors of Pivot collection, the simple (standalone) collection, the linked collection and the just-in-time (JIT) collection. The metadata for a simple collection is entirely defined in a single cxml

file and refers to a single deep zoom collection (.dzc) of images. This simple collection is limited to about 3000-5000 items. For larger datasets, the items can be split up across multiple collections which link to one another. Both of these types of collections have the advantage of being relatively simple to create and require no special logic on the server.

Besides being limited in size, however, they are also rather static. Changing the collection often means regenerating large portions or all of the material. For extremely large or dynamic datasets, software can be written that runs on the server and creates the data on-the-fly in response to the Pivot viewer's HTTP requests. APIs and SDKs are available to facilitate the more complicated logic required. This JIT method offers the ultimate flexibility at the cost of complexity and requires more technical ability to produce (see http://www.silverlight.net/learn/pivotviewer/collection-design/).

Tool for Creating Pivot Collections

There are three tools for creating pivot collections. First, an Excel add-on allows the user to assemble collection metadata in a spreadsheet, link to image files and press a button to output all the necessary files for the collection. This is often the easiest approach and requires no programming. Another easy approach is to use the free command-line tool. In this approach, the user structures the data in a well-known XML format, then invokes a command-line tool to create and link the images and metadata into a Pivot collection. This approach offers a little more flexibility as well as the ability to automate the process, but requires more technical knowledge.

The final option is to use one of the free APIs to write out Pivot from a custom program. This approach would allow the maximum amount of flexibility at the cost of requiring programming. For example, one could write a program that combined data from multiple databases or web services in an automated fashion to produce the collection. These tools are freely downloadable here: http://www.silverlight.net/learn/pivotviewer/collection-tools/

Genealogical Value of Pivot

We see several benefits to using Pivot to display genealogical data (especially thumbnails of scanned historical documents). First, many web users are unfamiliar with the "recall consequences" of multiple filters on a search query. Pivot's interface shows these consequences in a dramatically visual way that encourages users to experiment with different query criteria, discovering the characteristics of the collection as they go. Second, many collections of scanned genealogical documents have multiple axes of navigation (e.g., geographic, alphabetic, chronological), even within a single collection. In traditional, hierarchal browse structures it can be costly to adjust criteria in these multiple axes where much navigation up and down the tree structure is required.

Finally, Pivot encourages the use of visual cues to help patrons quickly notice outliers or key characteristics in a series of "small multiples." These cues can either be found in image content itself—examples include significant format changes in the historical image flow or interruption in the signal such as microfilm titleboards or blank pages. Such visual cues can even be designed by the publisher by surrounding Pivot images with a baseball-card-like border or other graphical elements, the colors or contents of which play an important semantic role in comparing and contrasting collection elements to each other.