Overview and Implementation Planning

Overview

What is the Variations System?

The Variations Digital Music Library System enables institutions such as college and university libraries and music schools to digitize audio and score materials from their own collections, provide those materials to their students and faculty for teaching, learning, and research use in an interactive online environment, and respect intellectual property rights. The Variations system offers users the ability to play sound recordings, view musical scores, create bookmarks and playlists, and create and save visual annotations for both audio and scores.

Variations is not a digital music *collection*. Instead, it is a system for institutions to provide their constituencies with access to institutionally owned or licensed content. Variations is most often used to provide controlled access to online music course reserves. But Variations is also useful for providing online access to sound or score collections generally. It is possible to use Variations as the front end for subscription databases such as the DRAM. Variations can also be used to provide online access to non-music audio such as broadcast or oral history collections.

Variations is a digital *library* system. As such,

- It is album-based, not track based: users see tracks in their album context.
- Score volumes are presented in their entirely.
- Access can be set up to Variations items from institutional OPACs, and links to individual OPAC records can be put in the audio player or score viewer.

The flexible access control system in Variations helps institutions limit access in accordance with local needs and legal interpretations.

- Variations does not use Digital Rights Management (DRM).
- Audio is streamed.
- Access can be limited to authenticated logins based on IP address and/or authorized group membership, such as a class roster.

Where did Variations come from?

The Variations system has been created as a result of a series of grant-funded projects at Indiana University. Funding agencies have included the National Science Foundation, the National Endowment for the Humanities, and the Institute of Museum and Library Services (IMLS). As part of the IMLS grant for the Variations3 project, Indiana University has released the Variations system as open-source and helped other institutions implement the system.

The original Variations project at IU went live in 1996. The technical infrastructure was completely reworked and a new system deployed in 2005, as a result of the Variations2 project. The IU implementation contains approximately 17,000 recordings and several hundred scores. It serves as the primary music reserves system for the Cook Music Library and the Jacobs School of Music. Every semester, students and faculty rely on Variations for thousands of reserve recordings and scores.

As a result of the Variations3 project, Variations is now in use at the the New England Conservatory, Ohio State University, the Philadelphia area Tri-College Consortium (Haverford, Swarthmore, and Bryn Mawr Colleges), and the University of Maryland. At least a dozen additional institutions are currently implementing or evaluating the Variations system.

What technologies is Variations based on?

Variations is an open-source system that relies on other open-source or freely available technologies, including the following.

- Linux. The server runs on the Linux operating system (or other Unix), typically RHEL 5.
- MySql 5 database.
- Java. Most of Variations is written in the Java programming language. The client application can run on Windows PCs or Mac OS X.
- Darwin streaming server. This is the free version of Apple's Quicktime streaming server, which can also be used.
- Djvulibre. The free version of Djvu for score image compression and viewing.
- Flash Player. The new browser-based audio player uses Flash for the user interface.
- Apache web server.
- Tomcat web application server.
- XML. Variations makes heavy internal use of XML and uses XML for users' data files.

How can other institutions implement Variations?

Variations is available as open source and is therefore freely available. Indiana University continues develop and maintain the Variations system and is investigating mechanisms to provide support. Meanwhile, see the SourceForge site for information on signing up for the support discussion list.

Implementation Planning

Leadership

A successful Variations implementation requires leadership from both information technology and library staff. As a technology-based system, Variations requires IT expertise to set up and administer, and typically needs to integrate with other information systems and processes. As a digital library system, Variations also needs to have leadership from the content owning organization such as the library or archive. Content owners need to establish intellectual property protection policies, manage access, develop priorities and workflow for content ingest, establish a preservation strategy, and consult with end users.

Planning a Variations implementation typically extends over a period of weeks or months because organizations may first set up a pilot instance to try it out. Even setting up a pilot instance may require ordering a server, setting up the server within the IT infrastructure, installing the third-party software, installing the Variations software, setting up a digitization station, loading some content, troubleshooting network and firewall issues, and so on. Full implementation may further require iteration with legal counsel on access policy, installation of the client software on classroom or library computers, distribution of client software to other users, creation of workflow documentation, customization of the Variations User Guide, etc.

Hardware Requirements

Server Hardware

What server hardware you use for Variations depends completely on how much content you plan to load and how much simultaneous use you expect the system to receive. Variations tends to be more demanding of memory and disk than of processor speed. Another variable is the chosen bitrate(s) at which you will encode. For planning purposes, assuming encoding at both 192kbps and 28kbps, a 74-minute CD will take up approximately 125 MB of disk space. The uncompressed wav file, if you choose to archive it, will require 650 MB of storage. Score pages might average 50KB each, compressed, depending on settings.

For network bandwidth, multiply 192kbps by the peak number of simultaneous streams you expect to figure out whether your network infrastructure is sufficient.

It is difficult to recommend a generic configuration for Variations. At the high end, you might spend $25k to $50, the latter if you want to have a backup system for automatic failover. But for most implementations a $8-10k system should be adequate, and a pilot could be run on a much cheaper or older system.

Digitization Hardware and Software

The Variations digitizer client only runs on Windows PCs, so a recent XP or Vista PC is recommended. Audio compression is CPU-intensive, so doing digitization on a slow system will render the system useless for anything else during compression. If analog audio will be put in Variations, a high-quality sound card or a USB (digital) connection to the analog input device is important. For score scanning, any flatbed scanner should work.

For audio capture and editing, commercial software products such as Sound Forge or WaveLab are typically used, although the free Audacity can also be used.

Planning and Staffing

IT Staffing Requirements

The fundamental need for system setup and ongoing administration is for a Linux system administrator familiar with downloading, installing and configuring free and open-source software. Specifically, the following components are involved in preparing a server for Variations:

- Linux (typically RHEL or equivalent)
- Network Alias Devices (ideally three fixed IPs)
- Firewall Configuration (opening ports for RMI, UDP, Apache, etc.)
- Java JDK
- Apache
- MySQL
- Perl (base Perl plus numerous additional modules)
- Darwin Streaming Server
- Tomcat (if deploying Variations web player or access manager)

Although we provide step-by-step instructions for setting up Variations, inevitably there will be institutional differences requiring good troubleshooting skills and ability to get help from other experts on campus. Extra help and consulting may be needed during the process by other IT staff such as network engineers (for networking configuration, firewall and bandwidth issues), authentication specialists, authorization system integration (if basing access on course rosters), storage management (for archiving uncompressed derivatives and backing up the server), and client support for Windows and Mac (if distributing client software to end users). Installing client software on end-user-managed systems will take additional support to help users troubleshoot installation problems. If the system will be integrated with a library catalog and/or z39.50 server, some library IT expertise may also need to be consulted.

After the server is prepared for Variations, the Variations system can be downloaded, installed and configured. The system contains a small amount of sample content that can be used to test the system after a digitizer's client is configured, built and installed.

Once Variations is up and working, it requires little involvement for ongoing operation beyond occasional trouble-shooting or upgrades (typically done only once or twice a year).

Key IT Decisions

Although the steps to set up a working system could be done in a day or two, initial setup and configuration often take longer due to coordination and integration issues. For this reason we recommend beginning the process at least two months before you want the system available for pilot testing.

Some of the key decisions to be made that will affect upfront IT time commitment include

- whether and how users will authenticate
- whether access control will be IP-based (easier) or enrollment-based (more work to tie in with student information system)
- whether to keep uncompressed audio and score image files and backups of the compressed derivatives, and whether archiving will be automated
- how to back up the database
- whether end-users will use the Variations client or just the browser-based web player
- whether end-users can install the Variations client on their own computers
- whether administrative users such as digitizers will be managed by IT or by the library
- how links to Variations content will be made available to users (e.g., integrating course reserves lists with a course management system could require some scripting)

Library Staffing Requirements

Library staffing for the ongoing operation of Variations is primarily for digitization activities. Digitization work can typically be done by student hourly employees as long as the workflow is defined and documented. Digitizers need basic skills in moving files around. Required audio skills depend on whether analog media will be digitized or not. For score scanning, skill level can be low as long as the workflow is carefully documented.

Putting a CD into Variations can take 10-20 minutes form the beginning to the end of the process. Creating compressed derivatives takes the most time, and other work can be done during the compression activity. If the track listing cannot be downloaded from the internet, entering the track titles can add some time, particularly for classical music where it may often be desirable to create a track hierarchy for multi-movement works.

For more information about putting content in Variations, see the Digitization and Cataloging Guide.

Key Library Decisions

During the initial planning and deployment of Variations, library staff will need to work closely with IT to ensure the system is implemented in a way that supports its intended use. Key decisions library staff are involved in include

- Access Policy. Who can view or listen to what content, where? Library staff typically develop an access policy (sometimes in consultation with legal counsel), but the policy needs to be reviewed and implemented by IT staff. An access policy may need to address student use, faculty use, source of content (library-owned or instructor-owned), recordings of an institutions's own performances, classroom use, distance education use, and textbook companion recordings. Here are some sample access policies in use (singly or in combination) at various institutions:
    - Anyone on the campus VPN can access any content.
    - Anyone on a computer in the music library can access any content.
    - Students on the roster of a class for which the instructor has requested Variations access can access any content.
    - Students on the roster of a class can only access content on the reserve list for that class.
    - Items not held by the library but belonging to faculty can only be accessed by students in a class for which those items are on reserve.
    - Music library staff can access any content.
    - Music faculty can access any content.
- Collection Policy. What materials will be put into Variations? Will instructor-made compilation recordings or score coursepaks be digitized? Will materials for large-enrollment classes have priority? What fragile-media materials could be better-accessed via Variations? What about recordings of institutional performances?
- Discovery. How will people find out what items are available in Variations? The Variations search window only works if extensive cataloging is done to the digitized content. Most institutions choose instead to make Variations content accessible from their OPAC and/or from reserve lists.
- Workflow. How will digitizers put content in Variations while tracking their work?
- Sound quality. The Quicktime streaming can handle automatic fallback from a higher bitrate to a lower-quality one if the network connection cannot support the higher-quality streaming. So some sites encode audio at two rates, 192kbps and 28kbps. Variations can handle other bitrates as well, but 192kbps seems to be a good compromise between bandwidth and sound quality.
- Preservation Strategy. The Variations digitization process typically starts with uncompressed wav and tiff files, but it does not deliver those files. Variations is not a preservation system per se, so each institution makes a decision about whether and how to provide long-term storage the compressed master files.