Archival Master – An RUcore Case Study

This case study is a summary of practices that Rutgers University Libraries has used in the treatment of archival masters to be deposited in RUcore (Rutgers Community Repository). These practices have developed over a period of years and are recognized as compromises between preservation theory and practice. As a result, the last section summarizes issues that might be valuable to others who are dealing with similar problems.

Preservation Policy

This section highlights the important policy aspects related to the archival master. Digital resources are part of RUL collections and subject to the same criteria for selection and retention decisions as other media. As such, they are included under the central RUL preservation policy: ensuring that the collections remain available over the long term through prevention of damage and deterioration; reversing damage where possible; and, when necessary, changing the format of materials to preserve their intellectual content. Critical technologies that are part of the RUL digital preservation program include:

- Persistent IDs for referential integrity using the CNRI Handle system
- Archival masters normalized to non-proprietary, open standards wherever possible
- Checksums on every archival master using SHA1. Checksums are periodically verified.
- Automated alerting to the Digital Curator if checksum failures are detected
- Documented object architectures and content models for generic classes of objects and verification of these architectures on ingest
- File format validation (prototyping work underway)
- Virus checking

Archival Master Defined

All generic object types (book, ETD, map, dataset, etc) have been defined to include an archival master. The primary objective of the archival master is to provide an uncompressed, non-proprietary, open data format that can be migrated forward as new standards and technologies become available. According to RUcore specifications, the archival master encapsulates, in a tar file, the master file or files to be preserved. For example, the master for a map might be digitized at 600 dpi as a tiff file. Further, the tiff image might be enhanced or modified to improve display quality, resulting in what we call the master derivative. Both of these files are encapsulated in the tar file to create the archival master. Presentation files are generated from the master derivative if it is available, otherwise the master is used. In general, the presentation files are web-friendly file formats that are down-sampled from the archival master in order to make them suitable for online access by end users. Other examples are included in the discussion below on content models.

Technical Metadata for Archival Masters

The RUcore ingest process requires that a signature or checksum, using sha1, be computed for each archival master. The RUcore ingest pipeline automatically creates this checksum along with the file size and stores this data in the technical metadata for the object. Note that RUcore uses the METS profile as a generic xml object structure and technical metadata is one of the administrative metadata sections in the METS profile. After an object has been ingested, a periodic procedure for verifying checksums is initiated via a unix cron process. Given the file size and resulting cpu requirements to verify archival masters, checksum verification is done nightly. At present, all checksums for every object are verified once a week and alerts are sent to the digital curator if there are any verification failures.

For the RUcore repository, the Fedora audit trail is enabled so that changes to any object are recorded in that object. In addition, the RUcore reporting service records all ingests, purges, edits, and signature failures. This data has been very useful in tracking anomalies in objects.

Additional efforts are made at the cataloging level to assess the condition and integrity of the archival masters through the collection of some 60+ data points – some of which are repeatable - in the Technical Metadata specification for digital objects. These user-input fields capture specific information about the creation of the digital object, including application software used, operating system platforms, and the specific digitization hardware used.

Content Models and File Formats

As part of our migration to Fedora 3.0, we have implemented some 15 content models (e.g. audio, book, map, video, etc). These content models are generic and await further definition to support enhanced services for custom dissemination and validation. Descriptions of archival masters for specific content models is included below:

Audio. Sound objects, whether they are created from digital sources or digitized from analog recordings, are submitted as an uncompressed WAV file. Bitrates can vary depending on the type of audio recording and the fidelity required to faithfully reproduce the content. We do however, recommend 44.1kHz sampling or better for the majority of our audio objects, with a bitrate of 192kbps or better. Through this WAV file, which is stored as an archival data stream in a TAR archive, an MP3 presentation file can be generated by our Workflow Management System (WMS), or a manually-generated, optimized MP3 can be uploaded with the archival master.

Book. Our Book Content Model is intended to serve a use case where books are scanned in as images, with each page being scanned into a TIFF file, text extracted into an OCR layer for searchability, and two presentation formats are derived from the TIF files. The CM presently consists of 5 data streams, which include the TIFFs as archival masters; a PDF and a Djvu file for public viewing; a structure map to assist in generating a digital "Table of Contents"; and an XML stream containing the searchable OCR data.

Specific preservation-grade standards have been written for the proper creation of scanned archival TIFF files. We specify a minimum resolution of 400 dots per inch (dpi) for bitonal black and white text scans, and a minimum 600 dpi for scans of pages requiring greyscale or color. Lossy compression of any kind is prohibited in the archival master, however lossless compression such as LZW, is permitted. TIFF files can also make use of lossless ZIP compression, but this mode is not widely supported by all imaging software, and is not recommended.

A great deal of our book scanning operations are outsourced to third party vendors who have appropriate book scanning equipment and can provide fast turnaround. These vendors in-turn generate each of the datastreams for book objects, which are then ingested. Alternately, there are cases where a book is scanned in-house using on-site scanning equipment. In the latter scenario, TIFFs are the only data stream generated and provided to RUcore. Using our Workflow Management System, and a custom-developed imaging pipeline, the TIFFs are used to generate the remaining data streams using added helper applications which recognize text and make necessary image format conversions.

Photograph. Still images can originate either as digital surrogates – photographs that originated as paper/film/slide analog formats and later scanned into a digital form – or as born-digital content, where the image originated from a digital camera. Minimum standards have been devised for the scanning of analog still images, and recommendations have been outlined as to the minimum requirements for digital cameras used in preservation-grade digital photography.

For digital surrogates, a minimum resolution of 600dpi has been established for color and grayscale images. However, an additional requirement known as the 3,000-pixel rule is also enforced. For each photograph scanned, at least one axis of the digital surrogate must yield a minimum width of 3,000 pixels. This is to ensure a minimum level of detail even in cases where a 600dpi scan alone would not normally be adequate, such as when digitizing slides or 35mm negatives. For such cases, scans as high as 2400dpi or greater may be necessary.

Digital camera requirements are based on the size of the imaging sensor, with a minimum 8 megapixel resolution required for archival use. Additionally, it is recommended that the camera be capable of storing photographs as TIFF files, or as RAW camera files. We recognize that compelling events of historic significance may be spontaneously captured on lower-resolution cameras. In such cases, and where an image of higher quality is unavailable, best efforts are made to preserve the available photos.

Regardless of origin, data streams consisting of our present delivery formats, djvu, pdf and jpg, are created by the workflow pipeline from the archival master TIFF file supplied during the cataloging process. When possible, born-digital photographs add an additional Digital Negative (DNG) file, which is a standardized method for encapsulating Camera Raw digital files. This permits us to capture and preserve not only the contents of the digital photograph, but the sensor telemetry and metadata that is generated and embedded by most modern digital cameras into the photographs they take. For the vast majority of photographic digital objects, text recognition is not performed, unless a significant amount of readable text is present.

Map. From a formats perspective, maps are treated in a matter quite similar to still photographs by RUcore. However, there is additional metadata that can be captured, such as bounding coordinates and relevant place names. File size becomes a greater concern as maps tend to be of a much larger physical format than a typical photograph, and requires specialized equipment and higher pixel counts to effectively capture the content. For the majority of map objects, an in-house large format scanning system, consisting of a 40-megapixel digital camera suspended over a flat scanning platform, provides the digital surrogates for scanned maps.

Video. (See object architecture in Figure 1). Video has posed the greatest challenge to digital preservationists, and has given us an opportunity to test the true adaptability of RUcore and the underlying Fedora platform. In particular, video objects were the first major object type we encountered with significant rights and encumberment issues. Previously, the vast majority of documents and photographs stored in RUcore were open and unrestricted, permitting us to deliver the content to all visitors without having to urgently explore or implement access controls based on rights restrictions. Video changed this, and required us to rapidly develop delivery methods that would make reasonable efforts to protect rights-restricted content.

Our greatest challenge, however, was that moving images have an inherent tendency to produce exceptionally large preservation data streams, larger than any other object type that RUcore had endeavored to ingest. Typically, an uncompressed, full-frame video file will take approximately 10 GB of disk space for each half hour of recorded Standard Definition video, with HD content far exceeding this estimate. This is problematic in system limitations restrict data streams to no larger than 2GB (see issues, below),

An early interim solution involving segmented archival data streams that could later be reconstituted to retrieve the archival master video file was tried, but proven to be impractical and risky. Fortunately, current versions of Fedora support external data streams, and we have implemented this architecture into the default content model for video.

The CM ultimately implemented consists of an externally managed Archival Datastream, which can take the form of an uncompressed AVI file for video converted from analog formats, or a capture of the digital master content for born-digital video (such as video captured in DVCAM, MPEG-2, AVC MPEG-4, etc.), bundled into a tar file. Depending on the nature and condition of the archival master, high-quality derivative “helper files” such as a high-bitrate H.264 video stream may also bundled into the tar archive to potentially assist future transcoding efforts, and act as a “sanity check” to compare against the original master.

From the archival master, presentation data streams are manually generated that balance the need for quality video playback while keeping in mind the constraints of some internet connections. In general, a broadband connection (at least 768 kbps or higher) is necessary to view the video content. We produce one high-bitrate (up to 1Mbps) Quicktime H.264 video stream, and a 512kbps, slightly-lower quality progressive-download flash video using ON2 VP6 compression. Additional datastreams include a Structure Map, which permits navigation of the video and provides a development point for our upcoming video annotation services, an XACML policy to define rights and access restrictions, and when available, a PDF data stream for transcripts.

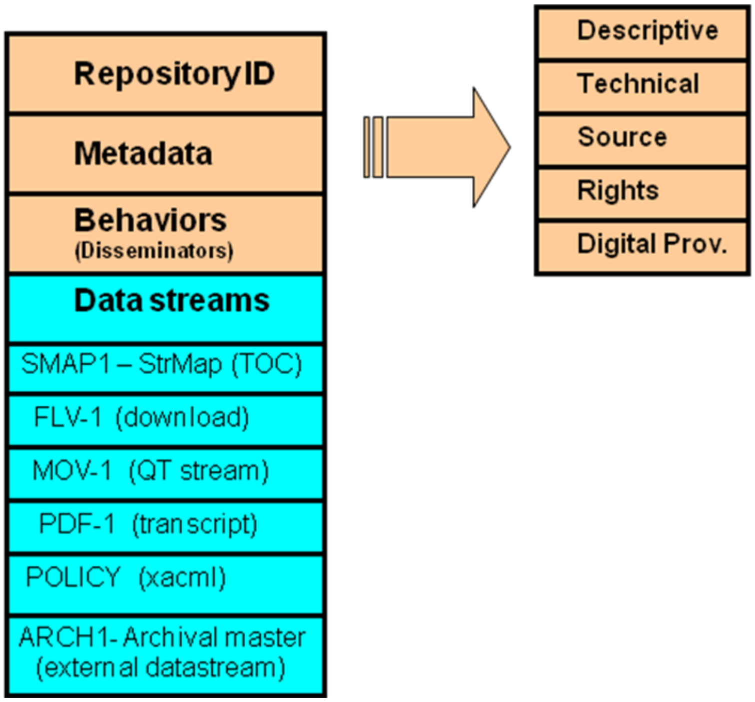

Figure 1 – A Video Object Illustrating Datastreams and METS Metadata Profile

Storage Backup

Past history as well as project team members' hands-on experiences with digital preservation have prepared us for the reality of implementing a robust backup strategy. A multi-tiered approach has been put in place to protect the objects stored by RUcore, and ensure the longevity of the data. Our approach entails the use of active, on-line data integrity protection, and the use of both on-site and regularly-rotated off-site copies of repository objects.

The primary storage system for RUcore is an expandable hard disk array, configured as a RAID 5 to enhance both the speed and reliability of the system. Currently, the storage system provides nearly 10 Terabytes of redundant online storage, but planning is underway to greatly increase its capacity. This system is our first line of defense in keeping objects safe and accessible. In any magnetic disc-based storage system, eventual media failure is a fact of operation, as is the possibility of data corruption as the stored objects age. The redundancy and built-in error checking attributes of the RAID permit RUcore to continue operating even as such a failure occurs, and provides ample opportunity for support staff to enact the necessary measures to recover from the failure mode.

For RUcore, digital objects enter the backup system within minutes of their being committed to Fedora. An automated tape system makes near-line incremental backups of these objects for efficient on-site access. This system is intended to provide us with a prompt method for the restoration of objects which fail a digital signature check, or to recover – if necessary – from a hardware or software-related event which results in data loss on the disk array. This tape backup system has been used once since RUcore's establishment to recover from such a loss; in January 2008, when a rare adverse reaction between the repository's RAID hardware and vendor firmware occurred. The event, triggered by a single disc drive failure in the array, would have normally been a manageable event resulting in no loss of downtime. However, the interaction caused a prolonged failure, and eventually required the recovery of corrupted data streams by recalling objects from the near-line storage system. The recovery was successful, and the hardware ultimately restored to an optimal and reliable working state. We have since worked with the hardware manufacturer to prevent a repeat of the incident, and subsequent disk failures, which are a fact of life in a multi-terabyte storage system, have been better managed without loss of data or uptime.

The near-line tape backup system is expanded through the use of off-site backups. A complete weekly backup is routinely committed to tape and transported off-site by a third party vendor, who manages the security of the tapes and their rotation. This additional backup step is intended to preserve the digital objects in the event of a catastrophic event adversely affecting the datacenter, preventing it or the near-line tape backup from adequately protecting the integrity of the repository.

The current backup scenario outlined is fairly straightforward, and at present, not unlike the conventional backup policies in effect at numerous businesses and institutions with an investment in their data. However, these are practices whose reliability and robustness have been proven. Even so, we continue to closely monitor various initiatives including cooperative online backup and dark archive solutions such as LOCKSS.

Issues

Versions. Archival masters are typically quite large files. A master for a video in AVI format can easily range up to 20GB or more. The question arises as to how we should handle versions after multiple migrations. In a simple scenario, suppose we migrate maps in tiff format to jpeg2000 and, much later, to jpeg3000. If we have Fedora versioning turned on, the resulting digital object would contain all three files – the tiff, jpeg2000, and jpeg3000. For a single map image, this may not create an issue regarding storage management, however retaining multiple archival masters for a video or other sizable objects is likely to consume mass storage at an unacceptable rate. If one can speculate that there are n migrations over many years, the repository versioning policy might be one in which the nth, n-1st, and the first master are all retained. All other archival master files are deleted.

Fedora "Managed" Datastreams. In addition to the versioning problem, large files cannot readily be ingested using the http protocol. In RUcore, we initially established a preservation policy in which all file datastreams are under Fedora management. This approach offers some simplicity in managing and preserving digital objects and worked well for relatively small file sizes (e.g. < 500 MB). However, when file sizes grow beyond 500 MB, ingest times become increasingly unacceptable, even in scenarios where the Fedora repository and the digital object to be ingested resided in file systems local to the same server. The Fedora development group has indicated that local file ingest is a priority. There has also been discussion of a hybrid mode referred to as "managed external". Given the need for additional Fedora development, RUcore has tentatively made the decision to locate all archival masters in an external storage area (the datastream is marked with an "E"). Although this avoids the ingest delay due to large files, these archival masters are no longer kept under Fedora management and a non-Fedora management policy must be put in place.

Level of Preservation. At present, RUcore does not distinguished various preservation levels and it is assumed that each object will be migrated forward. However, certain collections do require different levels of attention. In particular, "born digital" material does not have a corresponding physical artifact and loss of the digital object can result in a total loss of the resource content. As result, "born digital" objects might require a different preservation level in contrast to those objects that are digitized. In addition, we are beginning to examine how we might deposit science data in RUcore. Granting agencies or the researchers who developed the data may identify a useful "life span" for the data, suggesting that preservation attention is not needed after a specific time.

Archival Master Contents. The archival master does not include the metadata. From the discussion of backup policies above, the metadata along with the archival masters is backed up on tape and a copy is also kept off site. However, if we move to a storage architecture which retains multiple copies (e.g LOCKSS), the question arises as to what is copied. In the future, it may be advisable to encapuslate the metadata into the archival master. This issue is one of practicality; frequently there are minor updates to the metadata (e.g. for typos or minor corrections). If metadata is encapsulated in the master file, this file would have to be opened and updated each time the metadata changes.

File Format Validation. From an archival and preservation perspective, it is desirable, not only to capture the versions of files, but to also validate these versions with respect to repository policies. For example, one might want to either prohibit ingesting or flag old versions of Microsoft Word in order to be able to monitor these out of date formats. A preservation policy might permit ingest of non-standard formats but not commit to migrating this content forward. There is a community effort underway to provide a file format validation capability, newly renamed Unified Digital Formats Registry or UDFR, which is building on the prior efforts from PRONOM and GDFR. One could imagine validating all file formats that are represented in the archival master. Given real-time ingest constraints, this file validation process may have to be executed in a pre-ingest process.

Recovery from data integrity events. The current use of tape backups poses challenges in maintaining accessibility while recovery is in progress. In particular, while tape backups do ensure that we have reliable copies of our digital objects, they do not address the need for real-time availability of the stored object in the event our primary Fedora repository server were to fail. We continue to research mirroring and other methods to enable real-time redundancy, including the use of a read-only "failsafe" server which contains a copy of RUcore's digital content, but is not intended for editing existing objects or adding new ones. A read-only mode system has been assembled and tested, and found to be feasible as a method to keep digital content accessible to end users.

Submitted by: Isaiah Beard – Digital Data Curator – Rutgers University Libraries
Ron Jantz – Digital Library Architect – Rutgers University Libraries

Ib/rcj – 08/26/2009