

**New Skills for a Digital Era
Case Study 1**

**Changes in Acquisition:
A Guide to the Ingest of Electronic Records**

Kevin L. Glick
Yale University

Eliot Wilczek
Tufts University

This case study focuses on the function of *Acquisition*, defining it broadly to include appraisal and accessioning activities. The study works under a broad notion of technical skills to include what resources an archives must have as an administrative unit to undertake a trustworthy acquisition (Ingest) process. These resources include policies along with hardware and software in addition to the theoretical and technical knowledge that its staff must have.

The case study is based on an Ingest Guide developed by Yale University and Tufts University as part of its NHPRC electronic records research grant project, Fedora and the Preservation of University Electronic Records (2004-083). A draft of the Guide is available at: http://dca.tufts.edu/features/nhprc/reports/3_1_draftpublic3.pdf.

The Ingest Guide refers to ingest broadly, defining it as the entire process of moving records from a recordkeeping system to a preservation system. In this process, the Producer and Archive define the scope of records to be ingested. Then, they agree to the manner of the transfer, validation, and transformation. Only at this point will the two parties be ready to proceed with the actual activities. Each part of the Ingest Guide includes a narrative summary, a flowchart illustrating all of its steps, a description of each step, and the components, resources, products, and documentation necessary to undertake each step.

In using the Guide to undertake acquisitions, archivists will need skills that will enable them to undertake the ingest steps, as well as to create or modify the resources that support those steps.

The case study examines two acquisition scenarios:

- | | |
|--------------|--|
| Scenario One | A hypothetical scenario of a university archives that undertakes a recurring set of routine, semi-automated accessions according to the steps of the Ingest Guide. |
| Scenario Two | A real-life scenario of digital video records transferred to a university archives on a series of DVDs according to the steps of the Ingest Guide. |

The case study explores the skills needed for each scenario by outlining in a table the parts of the Ingest Guide and identifying the skills needed to undertake the steps and/or develop and use the necessary resources. This description of skills is not part of the Ingest Guide and we have made them solely for this colloquium.

Skills are grouped into two Categories:

Technical

Includes knowledge of file formats, recordkeeping systems, programming languages; systems, network, database administration; schema creation capabilities; familiarity with validation tools.

Archival

Includes theoretical and practical knowledge of archival appraisal, processing, description, and ethics; understanding of copyright law; use of metadata; and management and administrative capabilities.

Scenario One

After every meeting of the University Board of Trustees the Office of the Board of Trustees (the Producer) transfers a CD containing the working files of the full board and committee meetings to the Digital Collections and Archives (the Archive). These records

Scenario Two

Graduate education in Art and Architecture being particular strengths of University, the Archive actively collects in two different record units. The first record unit, "Material of the School of Architecture," concerning events and exhibitions, consists of posters,

include agendas, minutes of the previous meetings, various reports, and a video message from the University President. Trustees receive these documents from a secure website a few weeks before each meeting. The records on the CD are in MS Word, PDF, and QuickTime Movie file formats.

The Producer and Archive have agreed to a Submission Agreement that calls for the Producer to transfer a set of working files to the Archive shortly after each Board of Trustees meeting on CDs via hand delivery with the records in a particular type of file format. The Archive describes in the Agreement its planned validation and transformation activities. This is a serial Agreement that apply to all successive acquisitions of the working files, allowing both parties to skip Section A of the Ingest Guide, Negotiate Submission Agreement, and focus on the actions described in Section B of the Ingest Guide, Transfer and Validation. Therefore, Section A portion of the table below focuses on the initial establishment of the Submission Agreement.

programs, brochures, and photographic images [print and digital] documenting events and exhibits. The second record unit, “Art, architecture, and art history theses and projects,” consists of research papers, theses, and projects by students. When initially created, the archival materials created and accessioned in these units were almost entirely on paper or three-dimensional objects. Today the Producer submits records upon the completion of an exhibition or the completion of a particular student project. The Producer transfers records to the Archive often.

The Producer and Archive have agreed to an informal Submission Agreement that allows for complete academic freedom, but also leads to a highly variable and often unstructured accession process. The Producer does not necessarily conform to specific terms and conditions of transfer, transferring records without a set schedule, in file formats that may not be feasible for the Archive to preserve, and often not describing the accession sufficiently. This is not a serial agreement and a new Submission Agreement is periodically necessary.

The Producer normally transfers records to the Archive in batches of between 1 and 100 CDs or DVDs. These disks contain a wide range of record types (from flat text reports to three-dimensional models to digital still image and video) and a number of different file formats.

INGEST GUIDE	SCENARIO ONE	SCENARIO TWO
A Negotiate Submission Agreement		
A1 Establish Relationship	<ul style="list-style-type: none"> • Archival skills (<i>appraisal</i>) to establish relationship with Producer as appropriate archive for Producer’s records and develop and interpret <i>Records Authority Statement</i> that serves as warrant for Archive-Producer relationship. • Archival skills (<i>descriptive and metadata</i>) to create a <i>Producer Record</i> entry for Producers • Archival and (<i>descriptive and metadata</i>) Technical skills (<i>XML schema creation</i>) to create machine-readable syntax for <i>Producer Records</i>. 	
A2 Define Project	<ul style="list-style-type: none"> • Archival skills (<i>appraisal</i>) to agree on scope of acquisition with Producer. • Archival skills (<i>appraisal</i>) to start a preliminary <i>Survey Report</i>. • Archival skills (<i>appraisal</i>) to determine Producer has appropriate authority over records based on <i>Producer Record</i> and preliminary <i>Survey Report</i>. 	
A3 Collection Information and Assess Value of Records	<ul style="list-style-type: none"> • Archival (<i>appraisal</i>) and Technical skills (<i>evaluation of electronic recordkeeping systems</i>) to survey records, determining their essential elements, authenticity, appropriateness for accession. • Archival (<i>appraisal</i>) and Technical skills (<i>evaluation of electronic recordkeeping systems</i>) to create the <i>Survey Instrument</i>, <i>Recordkeeping Evaluation Tool</i>, and <i>Records Retention Policy</i> to make this assessment. • Archival skills (<i>authenticity requirements</i>) to analyze and judge the grounds for presuming the authenticity of the records, which is particularly important with these Trustee’s records. 	<ul style="list-style-type: none"> • Archival (<i>appraisal</i>) and Technical skills (<i>file formats</i>) to survey records spread over a number of disks and stored in a number of complex file formats, and collect information about the records necessary to make an appraisal. This particular Producer happens to be on the cutting edge of multimedia technology and the Archive would be required to determining their essential elements, authenticity, and appropriateness for accessioning. Also, the Archive must have the skill to understand how to represent to essential elements of form for each arcane format that may need to be converted just to place the files in storage or even to simply view them. Without this step the assessment of the feasibility of preservation cannot be completed.
A4 Assess Record Types	<ul style="list-style-type: none"> • Archival (<i>appraisal</i>) skills to determine if the files are not one of Archive’s established record types. • Archival (<i>descriptive and metadata</i>) skills to create a <i>Record Type Record</i> for any new record type. • Archival (<i>descriptive and metadata</i>) and Technical skills (<i>XML schema creation</i>) to create machine-readable syntax for the <i>Record Type Records</i>. 	
A5 Assess Formats	<ul style="list-style-type: none"> • Technical (<i>file formats</i>) skills to determine if some records are not in formats that meet the Archive’s <i>Formats Standards Policy</i>. • Archival (<i>appraisal</i>) and Technical (<i>file format</i>) skills to create the following <i>Transformation Plan</i>: Transform Word files to 	<ul style="list-style-type: none"> • Technical (<i>file formats</i>) skills to determine if some records are not in formats that meet the Archive’s <i>Formats Standards Policy</i>. In this scenario, the Archive must also be able to reassess as necessary the Archive’s <i>Formats Standards Policy</i>. Many formats might be considered ephemeral or experimental

	<p>PDFs; keep PDF files as PDFs because that is already a preservation format for the Archive; manage QuickTime files natively, creating a new preservation format.</p> <ul style="list-style-type: none"> • Technical (<i>file format and XML schema creation</i>) to adjust <i>Format Standards Policy</i> written in machine-readable syntax. 	<p>in nature, but are necessary to preserve the essential elements of some of these records. In such cases, the Archive must be able to weigh the difficulty and expense of one preservation strategy over another (perhaps many of these files should be stored in their native format in order to preserve their look and feel). This would require expertise in determining the digital preservation strategy for that particular format.</p> <ul style="list-style-type: none"> • Archival (<i>appraisal</i>) and Technical (<i>file format</i>) skills to create any necessary <i>Transformation Plan</i>. • Technical (<i>file format and XML schema creation</i>) skills to adjust <i>Format Standards Policy</i> written in machine-readable syntax.
A6 Assess Identifier Rules	<ul style="list-style-type: none"> • Archival (<i>appraisal</i>) skills to determine that neither scenario requires the Archive to preserve any <i>Producer Naming/Identification Scheme</i> and to determine the appropriate <i>Archive Naming/Identification Scheme</i> for the records. • Archival (<i>description and metadata</i>) and Technical (<i>XML schema creation</i>) skills to create <i>Archive Naming/Identification Scheme</i>. 	
A7 Assess Copyright	<ul style="list-style-type: none"> • Archival (<i>appraisal and copyright</i>) skills to determine that the University owns the copyright to the records and based on its <i>Copyright Policy</i> it does not need to acquire the copyrights or a license to them. • Archival (<i>copyright</i>) skills to create <i>Copyright Policy</i>. 	<ul style="list-style-type: none"> • Archival and Technical (<i>digital copyright</i>) skills to understand the copyright issues surrounding DVD movies, which are a significant portion of the records in this scenario. Copyright restrictions (from the DMCA for example) may prevent the Archive from fulfilling its <i>Format Standards Policy</i> or <i>Transformation Plan</i>, or prevent the Archive from “ripping” the DVDs in order to store the records on central storage.
A8 Assess Access Rights	<ul style="list-style-type: none"> • Archival (<i>appraisal, ethics, and management</i>) skills to determine based on its <i>Access Controls Policy</i>, that the records should be assigned to a <i>Record Security Profile</i> that restricts access to the records for an appropriate period of time. • Archival (<i>management</i>) and Technical (<i>XML schema creation, recordkeeping system, and network</i>) skills to create this new <i>Record Security Profile</i> and determines that the Preservation System can accommodate these access control needs. • Archival (<i>management and ethics</i>) skills to create <i>Access Controls Policy</i>. 	
A9 Assess Recordkeeping System	<ul style="list-style-type: none"> • Archival (<i>appraisal</i>) and Technical (<i>recordkeeping system</i>) skills to use its <i>Recordkeeping System Evaluation Tool</i> to evaluate the web environment where the Producer keeps the Trustee working files and its process for copying the working files to CDs. 	<ul style="list-style-type: none"> • Archival (<i>appraisal</i>) and Technical (<i>recordkeeping system</i>) skills to use its <i>Recordkeeping System Evaluation Tool</i> to evaluate the web environment where the Producer keeps the Trustee working files and its process for copying the working files to CDs.
	<ul style="list-style-type: none"> • Archival (<i>appraisal</i>) and Technical (<i>recordkeeping system</i>) skills to create a <i>Recordkeeping System Evaluation Tool</i>. 	
A10 Assess Feasibility	<ul style="list-style-type: none"> • Archival (<i>management</i>) and Technical (<i>systems, network, database administration</i>) skills to determine, based on its current <i>Preservation System Capabilities</i> report, that this acquisition is feasible. 	

A11 Finalize Submission Agreement	<ul style="list-style-type: none"> • Archival (<i>management</i>) skills to finalize and endorse the Submission Agreement. • Archival (<i>appraisal and management</i>) skills and Technical (<i>XML schema creation, network administration, and recordkeeping system</i>) skills to create metadata requirements, transfer procedures and schedule, validation procedures, and SIP creation procedures for the records in the acquisition. All of these procedures become part of the Submission Agreement. 	
B Transfer and Validation		
B1 Create and Transfer SIPs	<ul style="list-style-type: none"> • Archival (<i>metadata</i>) skills for the Producer to follow the Archive’s SIP Creation Procedure and create the SIP(s) and transfer them to the Archive for each recurring acquisition of the Trustee working files. This includes copying all appropriate working files onto a CD, clearly labeling each document according to a specified syntax, clearly labeling the CD as belonging to a particular Trustee’s meeting, and hand delivering the CD to the Archive within a few weeks of each Board meeting. 	<ul style="list-style-type: none"> • Archival (<i>management</i>) skills to act on behalf of the Producer if (as is the case in this scenario) they are unable or unwilling to produce SIP(s) according to the terms and conditions of transfer.
	<ul style="list-style-type: none"> • Technical (<i>programming languages, XML schema creation, and validation tools</i>) skills to create tools and the processes to package the SIP(s). For example, during SIP creation, the Producer or Archive will need tools to create baseline measurements for integrity checks that will be verified later on. 	
B2 Validate	<ul style="list-style-type: none"> • Archival (<i>appraisal</i>) and Technical (<i>file formats and validation tools</i>) skills to check the SIP of each recurring acquisition for viruses; to check the success or failure of the file transfer; to check that the files are well-formed, that the Producer is authorized to transfer the CD, that the SIP contains all of the necessary records components, and that the components all validate. • Technical (<i>programming languages, XML schema creation, and validation tools</i>) skills to create tools and the processes to do this validation work. 	
B3 Transform and Attach Metadata	<ul style="list-style-type: none"> • Technical (<i>file format, programming languages, and XML schema creation</i>) skills to transform the records that were slated for transformation in Part A5 (this may mean to convert files from their native format to a preservation format) and to automatically attach metadata it infers from the Submission Agreement. The Archive would need sufficient skills to maintain file format conversion tools for all of the formats listed in the <i>File Formats Standards Policy</i>. 	
B4 Formulate AIPs	<ul style="list-style-type: none"> • Technical (<i>XML schema creation, and file formats</i>) skills to turn the records into AIPs according to its <i>AIP Configuration Rules</i>. • Technical (<i>XML schema creation, file formats, and programming languages</i>) skills to create <i>AIP Configuration Rules</i> and semi-automated process to generate AIPs. 	
B5 Assess AIPs	<ul style="list-style-type: none"> • Archival (<i>appraisal and metadata</i>) skills to assess a sample of the AIPs to ensure that they have the records they are suppose to contain. 	
B6 Formally Accession	<ul style="list-style-type: none"> • Archival (<i>management</i>) and Technical (<i>system</i>) skills to deposit the AIPs into the appropriate Preservation System and formally notifies the Producer of the acquisition. This may require specific system administrative and/or programming skills necessary to maintain and/or connect to the Preservation System. 	

