

Appendix D: Technical Appendix

This technical appendix is included to provide an extended description and discussion of two important concepts: (1) metadata and (2) electronic (digital) archives.

1. Metadata:

What it is: Metadata (data about data) includes all the contextual, processing, and use information needed to identify and certify the scope, authenticity, and integrity of active or archival electronic information or records. Metadata can come from a variety of sources. It can be created automatically by a computer, supplied by a user, or inferred through a relationship to another document. Metadata is created, modified and disposed of at many points during the life of electronic information or records.¹

Some metadata, such as file dates and sizes, can easily be seen by users; other metadata can be hidden or embedded and unavailable to computer users who are not technically adept. Metadata is generally not reproduced in full form when a document is printed.

What it does: Metadata may connect to electronic information or records in a variety of ways. The electronic information or record may contain a reference to the metadata, or vice versa. For example, a hypertext document may contain a link to an index that provides information about its context. A folder or directory listing may contain a reference to the location where the content of the electronic document is found.

Why it may be important: Certain metadata is critical in information management and for ensuring effective retrieval and accountability in record-keeping. Metadata can certify the authenticity of the content of electronic documents, as well as establish the context of the content. Metadata can also identify and exploit the structural relationships that exist between and within electronic documents, such as versions and drafts. Metadata allows organizations to track the many layers of rights and reproduction information that exist for records and their multiple versions. Metadata may also document other legal or security requirements that have been imposed on records; for example, privacy concerns, privileged communications or work product, or proprietary interests.

Metadata's importance in searching: Searching capabilities can be significantly enhanced through the existence of rich, consistent metadata. Searching is generally used in records management to select and/or classify data. For example, proper searching can help with the assignment of electronic documents, files and messages into appropriate records management categories. Metadata such as dates, folder information, subject designations and other properties can help generate or validate classifications of the item. Metadata such as e-mail thread information can be used to help assure that related items are maintained in context and/or treated consistently. If

¹ Examples of metadata (for electronic document files) include: a file's name, a file's location (*e.g.*, directory structure or pathname), file format or file type, file size, file dates (*e.g.*, creation date, date of last data modification, date of last data access, date of last metadata modification), file permissions (*e.g.*, who has read the data, who can write to it, who can run it). Metadata can also include user-input attributes, such as e-mail subject and addressing, keywords, content description, business purpose, and retention codes and classifications, and the person responsible for the record's retention and disposition.

descriptive metadata are the same or can be mapped across different electronic repositories, metadata can also make it possible to search across multiple collections or to create virtual collections from materials that are distributed across repositories.

Metadata and records management: Metadata can also play a crucial role in record lifecycle management. Organizations can design systems that will allow users to input information regarding retention periods and automatically identify or dispose of obsolete records based on those retention periods.

Where it resides: Some metadata is held in structures separate from the core electronic information or record, such as directories, listings and indexes of the files or messages, but may still be regarded as an integral part of the electronic information or record for certain purposes. For example, e-mail messages may be stored with a variety of metadata that may not be viewed by the end-user in the standard setup of the program used to view messages. This metadata may provide important information about a message, such as message thread information that may provide context for the message and a variety of date/time settings. A database may contain metadata, such as the time of entry or modification, the identity of the record's creator, and other information. Document management systems, which are programs designed particularly to preserve tracking and identifying information about electronic documents, hold a great deal of metadata.

The forms it takes: Metadata may be different depending on how or when it is accessed or viewed. For example, when a message is transmitted through an e-mail system it carries with it a variety of metadata, such as the date of creation, transmission to the recipient, and receipt, and the identity of all recipients, including those sent blind carbon copies. After the message has been stored by the recipient, "bcc" information may no longer be directly available to him or her. Yet, when the message is stored by the recipient, "storage level" metadata, not available while the same message is in transmission, may become associated with it. Such storage level metadata may include the folder in which the message is stored and the dates and times it has been re-forwarded or replied to by the recipient.

Metadata migration: For records to remain accessible and intelligible over time, it may be necessary to preserve and migrate the metadata associated with those records. If records that are currently being created are to have a chance of surviving migrations through successive generations of computer hardware and software, or removal to entirely new delivery systems, they will need to have metadata that enables them to exist independently of the system that is currently being used to store and retrieve them. Technical, descriptive and preservation metadata that documents how a record was created and maintained, how it behaves and how it relates to other records will all be essential.

Metadata considerations: There will always be important tradeoffs between the costs of developing and managing metadata to meet current needs, and creating sufficient metadata that can be capitalized upon for future, often unanticipated uses. As organizations develop records systems, they should consider which aspects of metadata are essential for what they wish to achieve and how detailed they need each type of metadata to be. An organization may require frequent ad-hoc discovery searches across information systems, protection from inadvertent destruction of documents or e-mail messages, or it may need to prevent disclosure of sensitive trade secrets from being re-distributed or copied.

It should be noted that some software applications carry forward the original author's name in the metadata. Thus, if another person, in creating a new record (*e.g.*, a letter), copies it and then modifies it with new information, it may still reflect the name of the original creator of the record

used to recreate the format in the metadata of the new record. In such case, the metadata for the new record may be misleading as to the “real” author of the new record.

Metadata standards: National and international guidelines (such as DOD 5015.2, ISO 15849, Model Requirements for The Management of Electronic Records (MoReq), or ISO 23950 (formerly Z39.50) can be extremely helpful in making sure that an organization’s metadata standards meet the needs of the organization’s users.

Transmission of metadata: Individuals who create and transmit electronic documents are often unaware of the existence of readable metadata that may inadvertently reveal privileged or confidential information to adversaries and other outside parties. Organizations should consider adopting policies to provide guidance to users regarding the transmission of metadata. Moreover, many organizations publishing data on “nets” (extra, intra, inter) may not be fully aware of the metadata that may be indexed by outside search engines and viewed by individuals outside the organization.

There are a variety of methods for managing and controlling the extent of metadata transmitted with the core data. Some formats designed for transmission of data, such as XML, provide the functionality for the organization to determine which metadata fields are and are not transmitted with the core data. Other formats, such as the Adobe Portable Document Format (PDF) or Tagged Image Format (TIFF), can be used to remove certain metadata from the core document and to standardize the manner in which the document is maintained. Yet another approach is the use of “metadata stripper” technology, which removes some or all of the metadata from a native electronic file; however, such technology is not available for all types of data and may not be easily usable by end-users. Other technologies may be available for these purposes. Each technology embodies a different approach to the storage and transmission of the core document and metadata, and each may be appropriate in a given set of circumstances, depending on a variety of considerations, including usability of the data, cost, governmental rules and regulations, and other factors.

Metadata and new technology challenges: Emerging technologies may make the management of metadata in the electronic records context much more difficult. For example, “virtual foldering” may allow users to apply several different sets of metadata to a given electronic document depending on the context in which the document is viewed or processed. The metadata in this scenario may not be associated with a single document, but shared across a set of documents through a non-document information stores. As technology advances, metadata continues to evolve.

Some types of metadata continue to undergo changes that may increase the difficulty of electronic records management and production of electronic documents for legal proceedings. For example, on some (but not all) existing systems, the user or system administrator can control access to and usage of files and messages by rights or permissions. These constraints can themselves be important metadata properties for legal or records management purposes, and can also impact an organization’s ability to store or review its own data. In order to assure that all data can be accessed for purposes of the legal or records management function, permissions or rights to the data must be taken into consideration. Likewise, the legal and records management functions can be affected by encryption of data, procedures for compression and encoding, and other technologies that can make data difficult to identify or review.

One emerging technology that may have a significant impact is known as “electronic rights”, which refers to increased control over data access, storage and copying to prevent unauthorized use,

primarily in the copyright-protection area. Technologies designed to enforce electronic rights may cause records to be automatically soft-deleted prior to the expiration of its appropriate retention period, or may prevent the record from being reviewed or copied where necessary for records management or litigation purposes. Particularly in the area of audio-visual files (including voice mail and video recordings) the potential for restrictions in this area are significant.

2. Electronic (Digital) Archives:

What they are: Electronic archives are repositories for electronic records in a form that facilitates searching, reporting, analysis, production, preservation and disposition. When properly set up and maintained, electronic archives are not solely static collections of records (whether on-line or off-line on mass media such as tapes or optical media).

The importance of metadata in electronic archives: The key to maximizing the utility of an electronic archive is the availability of record metadata—especially metadata that cannot be easily derived from the record content—and record management data (such as the business owner, the planned disposition date, various retention factors, etc.) along with the native record. This additional data may add value for searching, reporting and analysis purposes. By adding value for business or user processes, electronic archive systems can present a positive situation for all parties within an organization.

Policies for access to long-term electronic archives should consider requirements for current and post-disposition access to metadata and statistical information.

Long-term business needs for metadata should be weighed against risk and record management requirements for comprehensive removal of both records and their associated metadata at the planned disposition point. These long-term needs may include compliance reporting, productivity analysis, project task and cost analysis, and other forms of detailed and statistical reporting.

Forms of electronic archives: Archives may be monolithic systems encompassing all functions required to create, retrieve, update, and delete electronic records across an organization, or they may be made up of multiple integrated electronic systems. This latter architecture is particularly appropriate for large organizations which already have document management (“DM”) or knowledge management (“KM”) systems in-place.

Integration of DM/KM and RM: The European Communities’ “Model Requirements for the Management of Electronic Records”² (“MoReq”) distinguishes between a DM and RM system (equivalent to an electronic archive in this context) as follows:

DM System ...	RM System ...
Allows documents to be modified and/or to exist in several versions.	Prevents records from being modified.
May allow documents to be deleted by their owners.	Prevents records from being deleted except in certain strictly controlled circumstances.
May include some retention controls.	Must include rigorous retention controls.

² Available at <http://www.cornwell.co.uk/moreq.html>.

DM System ...	RM System ...
May include a document storage structure, which may be under the control of users.	Must include a rigorous record arrangement structure (the classification scheme) which is maintained by the Administrator.
Is intended primarily to support day-to-day use of documents for business.	May support day-to-day working, but is also intended to provide a secure repository for meaningful business records.

Many DM/KM systems contain electronic archive (or electronic records management) functions, either as part of the base system, as add-on components or available through programmatic features. Where those functions do not exist for the system, it may be necessary to integrate stand-alone DM/KM and electronic archive systems by means of a real-time or periodic transfer between the respective repositories. The development effort involved in this integration can be significant. Both the MoReq and DoD 5015.2-STD³ provide useful starting points for defining integration requirements.

Electronic archives and e-mail: For most organizations, the ability of the electronic archive to work with existing e-mail systems will be critical. David Stephens notes:

... the management of e-mail is sometimes characterized as the single biggest records management problem in the USA. Thus, for any organization looking to implement major initiatives in the management of its electronic records, e-mail systems should be the initial focus of such efforts.⁴

Integration of e-mail can vary from simple journaling (also called “logging”) of all messages to the electronic archive, to interactive interfacing with the client e-mail application (for example, adding record classification functions to Microsoft Outlook). At a minimum, electronic archives should be able to serve as a repository for e-mail records exported from the e-mail servers. Many commercial e-mail archive and records management add-on products are available for popular e-mail systems (such as Microsoft Exchange and IBM/Lotus Notes).

Electronic archives and technology changes: As new applications are developed or acquired within organizations, the records management requirements relative to those applications should be anticipated and planned as part of the system development or purchase process. Digital preservation requires routine efforts to migrate records to overcome software and technological obsolescence and from deteriorating media.

Standards for electronic archives: Long-term electronic archive designs should consider incorporation of national or international specifications such MoReq or Open Archival Information System (OAIS). Standards such as ISO 15489⁵ establish guidelines for records management policies and systems but generally fall short of specifying functional details of automated systems. However, DoD 5015.2-STD MoReq contain useful information defining functional requirements for

³ Assistant Secretary of Defense (Command, Control, Communications and Intelligence) (2002), *Design Criteria Standard for Electronic Records Management Software Applications (DoD 5015.2-STD)*.

⁴ David Stephens and Roderick Wallace, *Electronic Records Retention: New Strategies for Data Life Cycle Management* (ARMA International 2003).

⁵ Available at <http://www.iso.org>. The two components of the standard are ISO 15489-1:2001 and ISO/TR 15489-2:2001.

electronic record archives. Both of these also define selected metadata elements required for an electronic records archive. Either document would be appropriate as a starting point for acquisition or construction of an electronic archive system. Finally, both ARMA International and the National Archives Records Administration (NARA) provide planning and guideline documents at their respective web sites.⁶

Tracking non-electronic records: Organizations designing comprehensive long-term electronic archives should consider the need for managing and tracking electronic and non-electronic records. This may include migration from legacy systems tracking paper, film/fiche, artifacts and electronic records.

Electronic archives and storage media: Policies for maintenance of long-term electronic archives should address selection of storage media and formats appropriate for data usage requirements and planned retention periods, including multi-format and multi-media transfers over the life of records. For the purposes of this discussion, “storage media” refers to the physical devices holding records. For electronic records this is typically fixed or removable hard disks, diskette cartridges (“floppy diskettes” of various sizes, high-density cartridge disks such as those manufactured by Iomega (“Zip disks” and “Jaz disks”) and Syquest), optical disks such as CDs and DVDs, or reel and cartridge tape. Excluding the optical disks, all these media store data electromagnetically and are capable of both reading and writing data through many “store-delete-write” cycles. Optical disks, as the name implies, store data by modifying the optical characteristics of a coated plastic disk. Some types of optical disks are capable of both reading and writing through many cycles; others are “Write Once, Read Many” (WORM)—meaning data can be written to the disk only once (that is, it is not updateable) but the disk can be read many times. The most common type of WORM disks are “CD-R” (“Compact Disk-Recordable”).

Storage media can be proprietary (controlled by a single corporation, often with details of the construction not available to other parties) or non-proprietary (typically controlled by a standards organization or a consortium of corporations; details of the construction may be available to other parties or restricted to members of the consortium). All present high-density cartridge disks and some forms of cartridge tapes are proprietary designs.

Significant issues may exist with media volume when used for archive purposes. At present, the highest density optical disks offer roughly 10% of the capacity of the highest density magnetic tape cartridges. Physical storage space requirements are comparable between the two (the amount of physical space required to store a given set of data) and storage arrays (“libraries” of multiple optical disks or cartridge tapes) exist for both media. Magnetic cartridge tape remains significantly more common for large-scale and long-term off-line and near-line storage in the corporate community.

When speaking of storage devices, the physical device is only half of the picture. The other half concerns how data records are stored on the physical device. “Format” refers to the binary representation of the data comprising a record. For electronic records there is usually a “native” format: the binary representation used by the application which normally creates, reads, and modifies the record as it is used during the active portion of its lifecycle. As an example, a project status report may be a Microsoft Excel spreadsheet; its format would be the proprietary binary format used by Microsoft for writing of this spreadsheet to storage media (informally this particular

⁶ Available at <http://www.arma.org>; available at <http://www.nara.gov>.

format is often called an “XLS file” because of the default file naming (“MyReport.XLS”, “Report701.XLS”, etc.) used by the Excel program). This format is called a proprietary format because its structure is “owned” and controlled by one corporation (Microsoft in this case). “Non-proprietary” formats may be public domain or made freely available for use by any organization. Some non-proprietary formats are nationally or internationally standardized. For example, the ASCII (American National Standard for Information Interchange) text representation coding is a North American standard. Others are de facto standards, an example of which is the PDF (Portable Document Format) binary representation for documents; this format is widely used by many Internet systems and document management applications).⁷

Ideally, long-term storage formats should be non-proprietary to avoid issues with technological and business obsolescence. However, in practice, non-proprietary formats may not support content and metadata information with sufficient fidelity to serve for archival purposes.

A well-designed electronic archive should support multiple storage media and provide mechanisms for tracking physical write date and time stamps for a given record (that is, the system should track when a record was stored on a given media—this is significantly different from the record creation metadata tracking when a record’s content was initially produced).

For records with long retention requirements it may be necessary to copy records to fresh media periodically. This process of copying to new media is referred to as “refreshing.” When should refresh copies be made? The National Library of Australia has concluded the best choices for long-term (over ten year) archival media and format are CD-R media and XML data formatting.⁸ Regarding optical media, they note “the lifetime of optical disks of all kinds, and especially CD-Rs, is greater than the technological obsolescence factor of their recording and playback technology.”⁹ NARA, in combination with the National Institute of Standards and Technology (NIST), provides guidance on CD and DVD media and formats in the NIST Special Publication 500-252, *Care and Handling of CDs and DVDs—A Guide for Librarians and Archivists* (NIST October 2003). The results of NIST’s evaluations are controversial and do not agree with manufacturer and independent testing.¹⁰ Given the significant variance among these expected life figures, a reasonable compromise may be to use the best quality media available, maintain both on-line and off-line media in an environmentally controlled space (stability appears more important than specific temperature and humidity values), and plan on refresh copies at intervals of no more than ten years.

⁷ PDF is copyrighted by Adobe Corporation but the specification has been made available for use by any party wanting to read or write documents using this format. Commercial applications writing this format may require a license from Adobe.

⁸ XML—Extensible Markup Language is a WWW (W3) Consortium standard; XML documents are encoded in UNICODE (itself an ISO standard for international character representations). Conceptually XML documents can contain any type of data (text, multimedia, numeric, etc.). In practice, XML documents are best suited for text and numeric information.

⁹ Ross Harvey, Presentation at the 2nd Nat’l Preservation Office Conference: Multimedia Preservation—Capturing the Rainbow in Brisbane (Nov. 28-30, 1995), available at <http://www.nla.gov.au/niac/meetings/npo95rh.html>.

¹⁰ A recent independent test on CD-R media concluded that many brands of inexpensive optical media have a useful life of less than two years. This contrasts dramatically with the NARA/NIST finding of an expected minimum useful life of 57 years. Refer to *PC-Active* (September 2003) for the most recent documented independent tests (available at <http://www.aktu.nl/pc-active/cdr.htm> (Dutch)); see *Development of a Testing Methodology to Predict Optical Disk Life Expectancy Values* (NIST 500-200), available at <http://palimpsest.stanford.edu/byorg/nara/nistsum.html>; last updated March 2002.

Due to rapid technological obsolescence, organizations may wish to consider duplicating particularly valuable records that must be kept for more than ten years to non-electronic media (*e.g.*, computer and output microfilm or “COM;” or archival paper).

Electronic archives and obsolescence: The electronic archive itself may be an application or set of applications. Over time these may change or become obsolete—often in less time than the longest retention period for the records associated with the system. For this reason, the archive architecture must anticipate and support future migration needs to new versions of the archive and the underlying storage media and formats.

Electronic archives and records destruction: Policies for maintenance of long-term electronic archives should address destruction and removal of records (and, as appropriate, their metadata) including any need for forensic-level electronic deletions. Methods for obtaining approval for destruction should be incorporated in the archive system.

Deletion of electronic records has a number of potential issues. In many electronic systems, there are two types of deletion: “logical” (or “soft”) deletions which mark record content as being unavailable (but do not immediately remove the record metadata or content) and “physical” deletions which remove a record’s content from its associated storage media (but do not necessarily remove all record metadata). Physical deletions typically require more time and computing resources than logical deletions. For this reason, physical deletions are often deprecated for systems requiring a high degree of user interactivity. Physical deletions may often be recovered; to prevent such recovery it is necessary to use a “wiping” technology that overwrites the deleted information in such a manner that it would require unusual (and expensive) techniques to accomplish recovery.

Deletion occurs in several levels on modern computer systems:

- (a) **File level deletion:** Deletion on the file level renders the file inaccessible to the operating system and normal application programs and marks the space occupied by the file’s directory entry and contents as free space, available to reuse for data storage.
- (b) **Record level deletion:** Deletion on the record level occurs when a data structure, like a database table, contains multiple records; deletion at this level renders the record inaccessible to the database management system (DBMS) and usually marks the space occupied by the record as available for reuse by the DBMS, although in some cases the space is never reused until the database is compacted. Record level deletion is also characteristic of many e-mail systems.
- (c) **Byte level deletion:** Deletion at the byte level occurs when text or other information is deleted from the file content (such as the deletion of text from a word processing file); such deletion may render the deleted data inaccessible to the application intended to be used in processing the file, but may not actually remove the data from the file’s content until a process such as compaction or rewriting of the file causes the deleted data to be overwritten.

Electronic archives should provide disposition functions for both logical and physical record deletions and permit specification of which, if any, associated metadata elements should be removed.

One issue that often arises is tracking details of when and how a given record may have been removed from the archive. In the paper world, “Certificates of Destruction” exist as proof that a set of records was destroyed by a particular method and by a specific organization on a given date. If a need exists for similar compliance documentation on electronic records, it will be necessary to keep a minimal set of metadata about those records to have a “target” for the data tracking the

disposition. This requirement will only exist if it is necessary to track the disposition information on specific records. Generic statistics (for example, a count of records deleted) can be maintained without retaining record metadata.

Electronic archives and security: Policies for access to long-term electronic archives should consider requirements for ownership and control including, but not limited to, security, traceability, and change-control over the record lifecycle.

The National Archives and Records Administration (NARA) *Concept of Operations*¹¹ provides useful guidelines for typical user functions and associated ownership concerns (references to “NARA” have been changed to “RM electronic archive”):

Access—All record users will be able to search and retrieve unclassified, unrestricted materials, which have been processed into the electronic records archive (ERA), either anonymously or by signing on as a registered user. Users with special access rights (clearances) and privileges will be checked for appropriate clearances by ERA upon accessing the system.

Search—The user searches ERA for information describing records and for actual content within records. Such searching may be done at a variety of levels of aggregation (documentary materials series/collections or individual items). Within the user’s given access rights and privileges, the user may take advantage of available functions and features. ERA responds to queries by identifying either sets of documentary materials, or individual documents, with results constrained by the user’s access rights. The user views and/or sorts the results of the search, modifies the search if necessary, and refines or saves query results as desired. In this manner, the user is able to progress from a query about a general topic to a list of specific documentary materials that the user may wish to view.

Retrieve/Receive—From search results that identify relevant documentary materials, the user views and accesses the records desired. The user directly interacts with the ERA system and accesses records in accordance with established user privileges and access rights.

User roles for electronic archives: When planning for specific control over the access, search, and retrieval rights of records in an archive there are a number of possible user roles. Users serving in these roles work in different ways—and at different times in the record lifecycle—with the archive itself, the record content and metadata, and the records policy infrastructure. Within the electronic archive there may be specific metadata associated with each role. The NARA *Concept of Operations* guide provides a working set of typical roles:¹²

Originating Entity (may also be called the “Author” or “User” in some contexts)—Creates and receives documentary materials and prepares and transfers them to the RM system.

¹¹ *Electronic Records Archives Concept of Operations*, § 6.6.2 (User Activities) available at http://www.archives.gov/electronic_records_archives/about_era/print_friendly.html?page=concept_of_operations_content.html&title=NARA%20%7C%20ERA%20@7C%20Concept. Note that this section defines additional classes of activities, specifically “Mediated Request” and “Fee for Service” functions, which do not apply in typical corporate archive environments.

¹² *Electronic Records Archives Concept of Operations*, § 5.3.1 (User Classes) available at http://www.archives.gov/electronic_records_archives/about_era/print_friendly.html?page=concept_of_operations_content.html&title=NARA%20%7C%20ERA%20@7C%20Concept.

Appraiser—Makes recommendations on materials that will be transferred to (RM system) holdings or will be disposed of by the Originating Entity.

Accession Processor—Accessions and processes a transfer (“accession” is the records management function of receiving a record or set of records into storage).

Preserver—Performs processing activities that ensure the ability to provide long-term access to documentary materials.

Access Reviewer—Reviews documentary materials in (the RM system) custody for access restrictions.

Record User—Uses the system to access documentary materials.

Administrative User—Handles such activities as granting user access rights, monitoring system performance, and scheduling reports.

This set should not be taken as absolute: many organizations will have only some of the roles, and some organizations will have additional roles. In particular, records management policies may define other roles (such as “Official Record Owner”, “Records Contact”, etc.) as appropriate for a given environment and organizational context. Finally, for electronic archives some roles, such as “Accession Processor” may be handled by automated agents (that is, by software rather than people).

There are additional Information Technology or Services (IT/IS) roles that may apply to an electronic archive system. These roles would be responsible for the creation and maintenance of the application software, hardware, and underlying database technology.

User management to control and track access, as well as change ownership and user roles, should be handled by an archive administration role. The NARA *Concept of Operations* guide refers to this role as the “administrative user” and describes three activities associated with the role:¹³

User rights and privileges—The administrative user assigns user rights and privileges based upon clearances held, permissions granted, job roles captured at the time of registration within the system, and RM policy.

Schedule Reports—The request for reports could be based on a specific requirement from RM policy or from a system monitoring need.

Monitor System—The Electronic Records Archive (ERA) provides the administrative user with the ability to monitor system performance and security.

The need for reporting functions: Reporting functions within the electronic archive—or the equivalent facility to report against the data technology underlying the archive (for example, to perform SQL (“Structured Query Language”) queries against an Oracle database on which the archive was built)—should provide access to historical, transactional and current record management metadata sufficient for auditing and verification of the archive. These tools provide the mechanisms critical to on-going validation of archive use, policy compliance, litigation analysis and extraction, and statutory or regulatory processing requirements.

¹³ *Electronic Records Archives Concept of Operations*, § 6.7 (Administrative User Scenario) available at http://www.archives.gov/electronic_records_archives/about_era/print_friendly.html?page=concept_of_operations_content.html&title=NARA%20%7C%20ERA%20@7C%20Concept.