Information Architecture

[page intentionally blank]

18 Data Management Architecture

18.1 Vision

The State of Kansas will provide data management services to ensure that digital information is accessible and reliable for employees, residents, and customers.

18.2 Scope & Business Rationale

State entities must protect critical data from loss or corruption in order to assure continuity of core business activities supported by that data. There are different ways to safeguard information including redundant storage arrays, parity checking, tape drives, optical storage, and storage management systems that can back up multiple storage arrays of varied type and manufacturer. Critical information must be backed up to a removable media, cataloged and stored in a secure manner that is appropriate for the specific data. This removable media may require off site storage and an archiving policy to further safeguard the State and the entity from loss or corruption of information. State entities need to match their specific needs and potential risks to the continuum of services and systems available.

18.3 Context & Diagrams

The data that resides on government computer systems is a vital state asset. Data has no value, however, if it is not accessible when it is needed. Effective data management requires an understanding of the systems that provide access to that data. Ensuring availability involves the following steps:

- Determining availability requirements
- Developing an availability plan
- Maintaining redundant data sources
- Managing long-term data retention
- Monitoring and reporting on availability

18.3.1 Determining Availability Requirements

Availability involves a number of factors, including the percentage of time that a service is available, the number of users who can access it and the speed with which they can access it. The first factor is determined over a fixed time period. A negotiated service time of eight hours a day with a 95% availability, for example, will allow for 24 minutes of down time during the service period. Such a figure must be approached with caution, however, since there is a significant difference between one 24-minute outage and 24 one-minute outages.

Availability requirements will vary from one case to another, depending upon the nature of the data and the needs of users. There are several dimensions involved in deciding the level of availability requirements of data:

Fluidity: How often does the data change? How often does the data need to be copied to respond to those accumulating changes?

Importance: How quickly must the original data be reconstructed to support business needs? How long must it be retained to support record keeping requirements?

Integrity: Must the data be copied in a specific way to ensure data integrity within the file? Across a group of files?

Point in time: Is there a specific time or interval when the data copy must be created?

Planning for availability is a continuous process, since it must account for changes in user requirements and technology. The Problem and Change Management processes can serve as early warning systems of the impact of change on the availability of IT service.

18.3.2 Developing an Availability Plan

Once availability requirements have been identified, an explicit plan should be developed to address the requirements. Service availability must be addressed in arrangements for service, which will often include service-level agreements (SLAs), for the protection of all parties involved. Planning for availability and developing meaningful plans is essential for ensuring that the processes are in place for meeting agency objectives.

18.3.3 Maintaining Redundant Data Sources

An important component of availability management is data redundancy. By strategically managing multiple copies of the same data, state entities can ensure the consistency, speed, integrity and quality of data access. Measures based on data redundancy include:

- Disk mirroring - Data written to one physical drive is also written to a second physical drive, sometimes called a hot backup, simultaneously. If one of the disk drives fails, the system can instantly switch to the other disk without any loss of data or service. This is often used in on-line database systems, when it is critical that the data be accessible at all times.

- Disk duplexing - Similar to disk mirroring, except that the second drive has its own controller. This is more expensive than mirroring but also more fault-tolerant.

- Server mirroring -All the processes and transactions of a primary server are duplicated on a second server, eliminating any down-time in the event of the primary servers failure.

- Data caching - A duplicate copy of data is stored closer to the user than the primary data source, in order to enhance performance.

- Clustering - Two or more computers are connected together in such a way that they behave like a single computer. Fault tolerance for a cluster is higher than for a single computer, but unlike mirroring, not all of the data is available on all computers at any one time.

- RAID (Redundant Array of Independent Disks) -- Disk drives employ two or more drives in combination for fault tolerance and performance. There are numerous RAID levels, each of which has different provisions for data storage. Level 0, for example, spreads out blocks of each file across multiple disks (called data striping), which improves performance but does not deliver fault tolerance. Higher levels provide fault tolerance through varying degrees of data redundancy.

- Routine system backups -- Backup and recovery of data is essential to any operation to ensure continuous data reliability in critical business functions. All data that could potentially require recovery is saved to offline media on a regular basis. The backup media generally have slower access speeds than hard drives and are less expensive to purchase and maintain. System backups should be stored in a different physical location from the primary source.

- Timed application backups -- Applications execute periodic point-in-time file backups to minimize the loss of current work in the case of premature termination. The backup files are generally overwritten with the next timed backup. Common examples are timed backups in word-processing and spreadsheet applications.

- Transaction logs - Changes to data are recorded and saved. This is most often used in database management systems, to allow rolling back to earlier versions of a data set in cases when that data set has been somehow compromised or corrupted.

No one of the measures listed above is sufficient by itself. Effective data management will make use of a combination. Each carries its own considerations related to retention and coordination.

18.3.4 Managing Long-Term Data Retention

State entities must often retain data for extended periods in order to ensure accountability for agency activities, meet legal requirements for records retention or preserve informational assets. In cases when the data has become inactive (i.e. it is accessed infrequently and users no longer alter it), it may be advisable to store the data on offline media, in order to save more valuable online resources and facilitate retention management. Unlike routine system backups, copies of the data on these long-term retention media will no longer be stored online. In this case, the offline version(s) is no longer a backup copy but is instead the master copy of the data. State entities may need to store multiple offline copies in order to meet business needs.

In order to ensure ongoing access to such historical data, state entities must:

- Maintain sufficient metadata and documentation to identify and apply appropriate retention periods to data (i.e. maintain accessibility of data that must be retained and destroy data that must no longer be retained).

- Refresh the physical storage media periodically to compensate for media degradation.

- Either store the supporting application software with data or convert it to new formats as systems change.

- Migrate data to new systems (desktop operating systems, network operating systems, enterprise management systems, etc) as they are implemented.

18.3.5 Monitoring and Reporting on Availability

The tools for monitoring availability are discussed in Section 2 on Network Architecture and Section 4 on System Management Architecture. The data collected by these tools is typically fed into a database, from which performance and trending information can be gathered in the form of availability reports. These reports should be used to evaluate system performance and identify areas that need improvement. In a case when a critical degree of IT services are not available, the situation could be considered a disaster; this topic is covered in more detail under the chapter titled Contingency Planning.

18.4 Principles

- State entities must have in place disaster recovery/contingency plans for mission critical data. The entity must ensure that such plans are available and that periodic testing is performed to assure proper functioning.

- Data back up and recovery procedures must be in place for mission critical data. State entities must have at least one type of back-up technology that is common to other state entities for recovery and data sharing purposes.

- Data management procedures must ensure that enterprise data is maintained in a manner that provides high availability, performance, and reliability to the end user.

- Data management architecture should be based on commonly accepted and/or emerging industry standards that are extensible, interoperable, and scalable.

- Data management procedures must address and ensure confidentiality of records that contain information that is limited for open/public access.

- Effective data management requires the creation and maintenance of metadata to document all information assets for back-up, recovery, and to ensure access for end users.

18.5 Goals

All state entities will develop, implement, monitor, update, and communicate with end users, when appropriate, the following data management components:

- policies and/or procedures for the effective protection, back-up, and recovery of all mission critical data resources.
- policies and/or procedures to ensure data integrity as required by Service Level Agreements (SLAs) and core business requirements.
- policies and/or procedures to ensure confidentiality of records that contain information that is limited for open/public access.
- data sharing/access policies and procedures that address core business requirements and related State statutes and policies.

18.6 Best Practices & Processes

- Information managers should work with business units/application/end user groups to define and implement strategic back-up/recovery policies and contingency plans to address mission critical data and end user requirements.
- A redundant hot site is recommended for state entities that collect and process data critical for the normal operation of any other federal, state, or local agency.
- State entities should maintain three generations of back-up media for daily processing and two generations of back up media for archived data. Back up media should be replaced on a periodic basis recommended for that media.
- State entities should develop regular reports that will provide pertinent information regarding data back-ups, their storage location, availability, retention, and expiration schedules.
- Within a state entity, or a consortium of entities, the control of back-ups should be managed by a centralized management system or operation.
- State entities should implement automated system management and monitoring methods of controlling backups and responses to failures. Key systems that ensure adequate data recovery and recovery response time should be tested periodically. End users share responsibility for testing backup/recovery functions to ensure proper operations.

18.7 General Standards

Category	Emerging Standard	Current Standard	Twilight Standard
Availability	Hierarchical Storage Management (HSM), Storage Area Networks (SAN), Clustering, High Performance Storage Systems (HPSS), Solid State Disk (SSD) Emulators	Redundant Array of Independent Disks (RAID), Data Caching, Server Mirroring, Disk Mirroring	Single Data Source, No Redundancy
Back-up and Recovery	Data Vaulting, Advanced Intelligent Tape (AIT), Virtual Tape, Digital Versatile Disk (DVD)	Compact Disk-Read Only Memory (CD-ROM), Digital Linear Tape (DLT)	Unix Dump, Floppy Disks, Desktop Backup Management
Data Archival	eXtensible Markup Language (XML) Wrappers, Records Management System (RMS) Software	Computer Output to Laser Disk (COLD), Computer Output to Microfilm (COM), Print to Paper, Offline Digital Storage Media	Floppy Disks

Table 18.7.1: General Data and Information Standards

18.8 Related Policies & Procedures

- Tampering with a Public Record (KSA 21-3821)
- Open Records Act (KSA 45-215?45-223)
- Government Records Preservation Act (KSA 45-401---KSA 45-413)
- Records made on Electronically-accessed Media; Authorization, Conditions and Procedures, Application, Notice to State Records Board (KSA 45-501)
- Public Records Act (KSA 75-3501---75-3518)
- Records Officer (KAR 53-4-1)

- General Records Retention and Disposition Schedule for State Agencies (KAR 53-3-1)
- Business Contingency Planning (ITEC Policy 3200)
- Business Contingency Planning Implementation (ITEC Policy 3210)
- Development of a Data Administration Program (ITEC Policy 8000)
- Kansas Electronic Records Management Guidelines (http://www.kshs.org/archives/ermguide.htm)

18.9 Technical Product & Configuration Information

Description	**Product**	**Notes**
Back-up/Recovery	Veritas Software NetBackup	Platform Independent
	Legato Systems	Platform independent
	IBM	Unix, NT, OS390
	Hewlett-Packard	Unix, NT
	EMC	Requires Symmetrix Storage for Full Exploitation
	CA ArcServe	NT, NetWare
	VM Center	OS390 Running VM
	Comm Vault	Strong Support for MS Exchange & Lotus Notes
	Storage Tek	
	Beta Systems/Harbor	
	Sterling/Spectralogic	
Storage Management	IBM Tivoli Storage Management (ADSM)	Sophisticated Core Technology, Supports Continuous Incremental Backup strategy at File Level-AIX, MVS
	HP OmniBack II	3-Tier, NT
	Legato NetWorker	Dominant in Unix, also NT, NetWare
	Veritas NetBackup	

Description	Product	Notes
	Interkink	MVS
	CA/Cheyenne	Small Scale NT, NetWare
	COMPAQ	StorageWorks, platform independent
	Spectralogic	Unix
	CommVault	

Table 18.9.1: Data and Information Product Information

18.10 Futures

Disaster Recovery and Backup for Operational Data

- Emerging technologies point to centralized Digital Linear Tape (DLT) Libraries as the most favorable storage management platform for rapid recovery of state agency mission critical data. DLT has a long history, but is designed to last. With a half-inch surface area and a tape length of up to 1,800 feet, DLT combines high capacity and long life. A single cartridge is good for 1 million passes across a tape head and can last for up to 30 years in storage.

- When Linear Tape-Open (LTO) and Super Digital Linear Tape (SDLT) products arrive early next year new solutions for large data centers that need fast data access and huge storage capacity will become available. These new tape technologies promise capacities equal to that of a mid-size state agencies entire financial database on a single tape cartridge. That could dramatically reduce the complexity of multiple cartridge backup systems, making it easier for state entities to have more effective backup and disaster recovery systems. To keep up with new developments in LTO and SDLT, visit www.lto-technology.com and www.dlttape.com.

- Advances in computerization require that state entities develop data archival management systems that include the archival of multi-media data and other complex data structures in their backup strategies. The restoration of those data types creates a challenge for state entities using them in their information systems. Increasingly, images, audio, and video are becoming a core component of state information systems. (eg. H/R database containing digitized images of state employees.)

- A more difficult challenge is emerging when attempting to record the complex relationships inside a ?canned? application used by state entities to manage records. (eg. Peoplesoft metadata layers between Oracle tables and end user reports.) A process to appropriately archive these types of complex data structures must be defined and clearly articulated.

- Guidelines must be developed to define an appropriate census point at which information kept in state information systems becomes a state record. Disaster recovery of a mission critical database may not constitute a state record, but merely temporary storage of information to perform state business on any given day. If that information constitutes a ?fluid? state record then we must articulate a census point where the data becomes a static state record.

Archiving for Official Record Purposes

- State entities must develop and implement strategies for data preservation that consider the projected life of the physical storage media, and the hardware and software used to write and read data that constitute a State of Kansas record. These strategies should provide for not only the refresh of physical media, but also, the conversion and migration of data to new formats and systems, as necessary.

- Data storage by state entities should be based on open standards that have been developed and approved by recognized industry standards bodies. This will reduce the cost of maintaining data stored in obsolete formats.

- Systematization of records preservation should include a clear articulation of the data and metadata elements that constitute official state records and the ability to read and rewrite those records to current storage technologies for long-term record keeping. State entities should also apply retention schedules to the records, so they focus their preservation efforts on records that warrant long-term retention.

- In order to allow for future access to records, it is critical to capture and maintain appropriate metadata along with the data used to conduct state business. This metadata will ensure the integrity of the content, context, and structure of the data as it was originally created, and will also convey vital system information necessary for future access to records.

- Two long-term storage technologies are on the horizon; HD-Rosetta, developed at Los Alamos National Laboratories, and HD-ROM developed by Norsam Technologies, Inc. These systems provide for the storage and protection of state records

using sophisticated technologies that are nearly indestructible and extremely compact. More information on these technologies can be seen at http://www.norsam.com/rosetta.html and http://www.norsam.com/rom.html.

- Other emerging approaches to long-term storage are represented by the Persistent Archives and Electronic Records Management project of the National Archives and Records Administration (NARA) and the San Diego Supercomputer Center (SDSC), which can be found at http://www.sdsc.edu/NARA/. They make use of eXtensible Markup Language (XML) wrappers to serve as persistent objects for records. This approach does not eliminate the need for ongoing work to maintain electronic records, but it does minimize it.

18.11 Organization & Personnel Impact

Due to advances in automation, robotics, and storage media in the areas of backup and archiving, training of systems personnel must keep up with advances in technology. This will be a consistent cost to organizations. In addition, some jobs may become obsolete as media and technologies change and personnel resources could be reallocated to other developing IT areas.

Each State agency must clearly articulate to all staff involved with maintaining and processing data that data are a State resource and must be appropriately safe guarded as outlined in this chapter and/or document. This is particularly true for data that supports official reporting to, and for, the State of Kansas.

In this age of high-powered computers on individual desktops, it is relatively easy for important agency databases to be created and maintained on an individual's workstation. For units that maintain mission-critical data in such a manner, the State agency should provide a centralized system or service that follows the guidelines found in this chapter. This resource could be as simple as making sure that the data are saved on a LAN at regular intervals and LAN administrators are following the appropriate archiving and backup guidelines.

[page intentionally blank]

19 Information Management Architecture

19.1 Vision

State entities at all levels of government and the public will have efficient, effective, and convenient access to accurate and current government information, as appropriate, under laws and policies governing security, privacy, and freedom of information.

19.2 Scope & Business Rationale

Information management architecture provides a framework for accessing data from online transaction processing (OLTP) systems and transforming data into various types of online analytical processing (OLAP) systems that help support Kansas agencies business decisions and which meet the informational needs of agency constituents.

Important data is stored in multiple applications within and across Kansas agencies. By grouping together this distributed data in a meaningful format, it can provide valuable information within an agency, as well as across agencies. Through the effective compilation of data, information can be created for reporting, records management, analysis, and decision support. These information resources can range from a simple database used to respond to frequent questions to a fully implemented data warehouse and/or clearinghouse. The goals for this information architecture are to encourage the development of appropriate and effective technologies to support these informational needs. The management of information should:

- Insulate transaction processing systems from the often large, ad hoc queries that are required by analytical processing systems.
- Provide a cross-organizational view of data where possible.
- Provide access to data not found in transaction systems, including summary data, historical data, metadata, and external data.
- Avoid the duplication of efforts to collect, verify, store, and maintain data used by multiple reporting systems and/or agencies.

- Provide an appropriate metadata repository that contains all the information about the data and processes used to populate and access a data warehouse and/or clearinghouse.

- Provide better end-user access.

19.3 Context & Diagrams

Information Architecture technology components assist in the management of data sources, the organization and definitions of data (metadata) and the extraction, storage, manipulation and retrieval of data contained in data warehouses, marts and other less structured storage formats. Components included under information architecture include the functions and items discussed below.

Business intelligence tools are used by end users for decision making and analysis. They allow the user to dynamically query the data and information stored in data warehouses.

Data Administrator: According to Information Technology Executive Council (ITEC) Policy 8000:

The agency Data Administrator, or a designated representative, shall participate in interagency data administration activities organized by the central data repository staff within the Division of Information Systems and Communication (DISC) with the assistance of the Data Sharing Committee. The Data Sharing Committee, in order to identify statewide Data Administration issues, shall make recommendations to the ITAB concerning, but not limited to:

- Standards relating to data as an asset to the State of Kansas;

- Data that are critical to the mission of the State, or common to multiple agencies;

- Policies that ensure the establishment of a statewide enterprise view of information;

- Enhancements to the state Data Administration Program;

- Minimum requirements for Agency Data Administration Programs; and

- Data administration education and awareness.

Data cleansers are used to validate and clean data so that the data is as consistent and accurate as possible, usually as part of a middleware implementation.

Data extraction and transformation tools (data movers) are used to extract and reformat data from legacy and OLTP systems according to metadata definitions. These tools put the data into a data warehouse.

Data marts are repositories of data gathered from operational data and other sources that is designed to serve a particular community of knowledge workers. The design of a data mart tends to start from an analysis of user needs, as opposed to a data warehouse, which tends to begin with an analysis of what data already exists and how it can be collected in such a way that the data can later be used.

Data mining includes the analysis of data to identify relationships that have not previously been discovered. Data warehouses become increasingly valuable as data mining approaches improve.

Data replication tools are used to distribute data from a data warehouse to various other data warehouses and data marts throughout the organization.

Data visualization tools are used to provide a wide variety of displays of data contained in the data warehouse (graphs, pie charts, maps, etc.) to assist in determining trends or patterns in interrelated data. The front end to a GIS system is a common example of a data visualization tool.

Data warehouses are central repositories for all or significant parts of the data that an enterprise's various business systems collect in order to represent a coherent picture of business conditions at a given point in time. Data from a diverse set of sources is selectively extracted and organized on the data warehouse database for use by analytical applications and user queries. Data mining and decision support systems (DSS) often make use of a data warehouse.

Decision support systems (DSS) are applications that are used to analyze business data and present it so users can make business decisions more easily. A DSS may present information graphically and may include an expert system or artificial intelligence (AI).

Enterprise resource planning (ERP) is a type of software that helps manage the important resources of an enterprise, such as product planning, purchasing, maintaining inventories, interacting with suppliers, providing customer service, order tracking, finance and human resources.

Geographic Information Systems (GIS) allow users to envision the spatial components of a data set by querying or analyzing a database and receiving the results in the form of a map. Information is organized in terms of spatial location such as geographic coordinates, street address, postal code, or taxing unit. Though they have been traditionally seen as separate systems, GIS data and applications are increasingly integrated with other information technology systems,

such as ERP, DSS, workflow and document management systems. GIS technology has evolved into a mainstream information technology through integration with relational database management systems (RDBMS). GIS has the capability to integrate various data from multiple data systems by organizing the data around its spatial location.

Government Information Locator Service (GILS) is a system used to identify, locate, and describe publicly available state government information resources, including electronic information resources. Though it can often be integrated with such systems, a GILS does not itself provide a user to directly access or manipulate the information resources themselves. A metadata repository, if based on sufficient standards, can serve as the foundation for a GILS.

Metadata is a definition or description of data, i.e. data about data. Metadata is used in the administration, management, discovery, retrieval, analysis, transfer, and preservation of data. It takes such forms as file names, directory structures, data dictionaries, indexes, catalogs, document type definitions (DTDs), structured tags and external documentation. From an agency perspective, metadata is essential for making sense of diverse data sets and appropriately dealing with things like rights management, security, and preservation. From the user's perspective, metadata relieves them of having to have full advance knowledge of the existence and characteristics of an object or data set. Metadata can be stored either embedded in the digital object it describes or as part of a separate but associated file. It is usually defined in terms of name/value pairs in which the name identifies the role of the specific element of metadata and the associated value indicates the term to be used to reference documents or digital objects exhibiting the required characteristic.

As indicated in Chapter 18, in order to allow for future access to records, it is critical to capture and maintain appropriate metadata along with the data used to conduct state business. This metadata will ensure the integrity of the content, context and structure of the data as it was originally created, and will also convey vital system information necessary for future access to records.

Metadata repositories are centralized collections of metadata, which should contain all the information about the data and processes used to populate and access a data warehouse.

Online analytical processing (OLAP) allows users to selectively extract and view data from different points-of-view. To facilitate this kind of analysis, OLAP data is stored in a database, which considers each data attribute as a separate dimension. OLAP software can then locate the intersection of dimensions and display them.

Online transaction processing (OLTP) is a class of program that facilitates and manages transaction-oriented applications, typically for data entry and retrieval transactions.

Schemas are documents that define the vocabulary used by an application.

Vocabulary is a set of metadata fields related to a particular application.

There are currently only a limited number of State of Kansas data warehouses and/or clearinghouses. The Kansas GIS Policy Board's Data Access and Support Center (DASC) is a functioning data and information clearinghouse with complete metadata, that organizes, archives, distributes and supports end users of spatial, or geographically referenced, data assets of statewide value. Additionally, both the KBI-developed CJIS system and the KDHE's "core database" (used to support overlapping federal and state reporting requirements) are centralized efforts to provide the beginning of intra- and inter-agency standards regarding data definitions, sources and uses, and end user access interfaces.

There has been considerable discussion about the creation of a statewide government information locator service (GILS) for information discovery. A common example of information discovery is a library catalog. It lets you know what books are there, but you generally cannot do a full-text search over the contents of the books themselves.

Such a repository and access system for high-level metadata would help address the needs of state customers and support compliance with the Kansas Open Records Act, KSA 45-221(a)(16), which states that each public agency shall maintain a register, open to the public, that describes:

- The information which the agency maintains on computer facilities; and

- The form in which the information can be made available using existing computer programs.

If this register requirement is addressed in a unified way across state government (consistent with the spirit of the KSTA and the SIM Plan), then the state must work towards a standard for information discovery and metadata documentation that addresses the issue of data content and data values. That is not only to indicate what metadata elements are included, but also what the content of those elements should be, and in some cases, even the specific set of allowable values. Compliance with such a standard for high-level metadata is necessary to make full use of the state's data and information assets.

The degree of control over the content of a given element will vary greatly, depending on what we determine to be both practical from the agency end and desirable from the user's end. In some cases, simply specifying data type and field length (if applicable) will suffice. But in others much more standardization would be desirable.

The overarching question is what minimal set of elements and constraints on the contents of those elements will be both necessary and sufficient for the diverse body of information held by the state. If we are to be realistic about a high-level information discovery metadata standard, it will have to be limited to only a core set of elements.

There are several standards developed to address this sort of heterogeneous search problem, specifically for online access to government information. They include:

- Global Information Locator Service (GILS)
  http://www.gils.net/

- Australian Government Locator Service
  http://www.naa.gov.au/govserv/agls/

These, and the majority of other such efforts for distributed heterogeneous information discovery use the Dublin Core (DC) (http://purl.oclc.org/dc/) as their foundation. The DC is a set of 15 elements that reflect a core of metadata about a resource (creator, title, date, etc.). One of the advantages of DC is that it is both basic and very extensible. Additional metadata for more specific data sets can be used by adding sub-elements, qualifiers or adding additional elements from other schemes, through a mechanism such as Resource Description Framework (RDF) (http://www.w3.org/RDF/).

The DC development community is also working on standards for the contents of each element, which implementers can either adopt or develop their own. The DC site indicates a huge range of projects, including government efforts that are making use of the DC. Two interesting uses of it by state government are:

- North Star, general government information locator
  http://www.state.mn.us/

- Bridges, Gateway for Environmental Information
  http://bridges.state.mn.us/

The Minnesota Metadata Guidelines - Dublin Core (MMG - DC) (http://bridges.state.mn.us/metadata.html/) contains detailed guidance, including a User Guide that provides assistance with the assignment of element values.

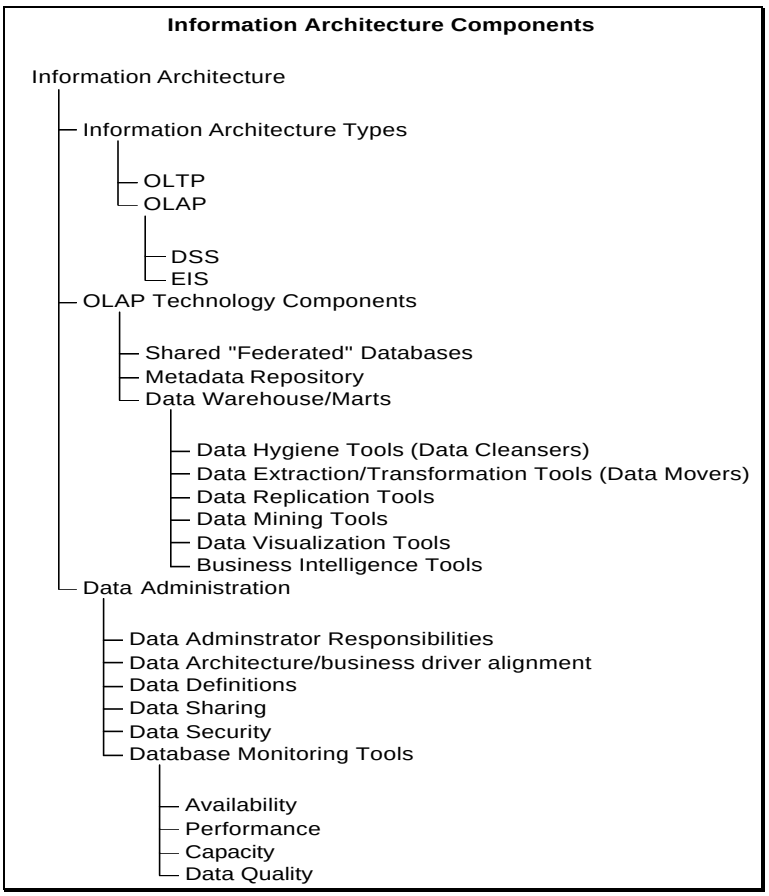

Figure 19.3-1 : Information Architecture Components

19.4 Principles

- State entities must seek to continually improve the quality, accuracy, and integrity of enterprise information through the promotion of data consistency and standardization.

- State entities must continually strive to improve data management and access through the use of appropriate existing and new methods, tools and technologies.

- The business functions and initiatives of state entities shall shape and drive the conceptual, logical and physical models of data and information assets.

- State entities must facilitate data and information sharing within the organization and with external user groups.

- Data and information resources are state assets that must be managed as valuable state resources, held in trust for the public.

- Data and information management practices and policies pertain to the entire lifecycle of the asset including its creation, use, storage, documentation (metadata), and disposition or archival.

19.5 Goals

All state entities will develop, implement, monitor, update, and communicate with end users, when appropriate, the following information management components:

- Kansas citizens will have access to necessary government information and services when and where it is needed, regardless of the geographical location.

- Inter-organization data sharing will allow redundant data stores to be kept to the minimum necessary for system performance, integrity, and security.

- State entity senior management and decision makers will understand the principles of sound data and information management.

- Overall management of data and information resources will be strategically aligned with the organizational goals and business practices of the state entity and enterprise.

- Metadata will be maintained for all data and information assets.

- State entities will encourage the use of common techniques and open standards to promote interoperability among systems and will identify opportunities for sharing commonly used data through integrated applications and databases.

19.6 Best Practices & Processes

- In the context of online analytical processing systems (OLAP), information lifecycle considerations are of paramount importance. Metadata should clearly delineate temporal aspects related to data validity and refreshment schedules.

- Exercise system lifecycle disciplines when planning the development, implementation, and maintenance of OLAP systems. Employ scaleable tools and design and build systems incrementally.

- Targeted business area requirements included in decision support systems (DSS) and executive information systems (EIS) must integrate with enterprise business requirements.

- Enterprise business drivers should define the frequency of database population and refresh methods.

- Adhere to project management procedures to measure dimensions required of OLAP systems. Performance tuning is a critical factor.

- When developing a data warehouse, start with a small, scalable system and then build incrementally.

- Adhere to project management procedures to identify and continuously measure performance along dimensions required of systems.

- Ensure the protection of individual privacy through appropriate security on confidential data.

- Ensuring the integrity and quality of data through quality control and auditing is the responsibility of both the business users and IS staff, particularly in the case of federated or warehoused data.

- Build data quality into new and existing systems.

- Design implementation plans to match business needs.

- Plan and budget for ongoing support and maintenance.

19.7 General Standards

The State of Kansas encourages the continued and expanded development of statewide standards for information architecture. The need for these standards will become increasingly more important as more mission critical inter-agency data and workflow sharing systems are deployed. Currently there are two converging standards for meta data, the first from the Meta Data Coalition (MDC) and the second from the Object Management Group. The two groups are working together on developing a Unified Modeling Language (UML). UML is a language for specifying, constructing, visualizing, and documenting the artifacts of a software-intensive system.

In relation to geo-spatial, or geographic information systems (GIS) data and information standards, the Kansas GIS Policy Board has adopted various Metadata and thematic data standards. Additionally, the Kansas GIS Policy Board as a cooperative partner of the Federal

Geographic Data Committee (FGDC), supports the development of the National Spatial Data Infrastructure (NSDI). The FGDC has developed, and/or is developing, numerous thematic geo-spatial data standards to guide the development of the NSDI. FGDC standards listed in the current column have been endorsed by the Kansas GIS Policy Board and the statewide GIS community. Standards listed in the emerging column are either in development or have yet to be endorsed by the Kansas GIS Policy Board. Access geo-spatial data standards at the following URLs: http://www.fgdc.gov and http://gisdasc.kgs.ukans.edu.

Categories	Emerging Standard	Current Standard	Twilight Standard
• Metadata • Analysis and design • Database and warehousing • Object and component design • Knowledge management • Business engineering	The Open Information Model Meta Data specification: • UML as a base model. • XML for metadata interchange. • SQL for data retrieval. (Refer to Meta Data Coalition at http://www.mdcinfo.com/)		
• Geo-spatial (GIS) Metadata		• ITEC Policy 5100, Kansas GIS Metadata Std (FGDC-STD-001-1998 V2.0, CSDGSM)	
• Geo-spatial (GIS) Thematic Data	• FGDC Classification of Wetlands and Deep Water Habitats • FGDC Vegetation Classification Std • FGDC Soils Geographic Data Std • FGDC Std for Remote Sensing Swath Data • FGDC Content Standard for Digital Geo-spatial Metadata, Part1: Biological Data Profile • FGDC Utilities Geo-spatial Data Content Std • FGDC Spatial Data Content Std, Computer-Aided Design and Drafting Profile	• Kansas Geodata Compatibility Guidelines V. 2.2 • Kansas GIS Cadastral Std • Kansas GIS Addressing Std • Kansas GIS Hydrography Std • Kansas GIS Administrative Boundaries Std • FGDC Geo-spatial Positioning Accuracy Std, Part 1: Reporting Methodology • FGDC Geo-spatial Positioning Accuracy Std, Part 2: Standards for Geodetic Networks	

Categories	Emerging Standard	Current Standard	Twilight Standard
	• FGDC Geo-spatial Positioning Accuracy Std, Part 4: Architecture, Engineering, Construction, and Facilities Management • FGDC Content Std for Framework Land Elevation Data	• FGDC Geo-spatial Positioning Accuracy Std, Part 3, National Std for Spatial Data Accuracy • FGDC Content Std for Digital Orthoimagery • FGDC Spatial Data Transfer Std • FGDC Spatial Data Transfer Std, Part 5, Raster Profile • FGDC Spatial Data Transfer Std, Part 6, Point Profile	

Table 19.7-1 : General Data and Information Standards

19.8 Related Policies and Procedures

- Development of a Data Administration Program (ITEC Policy 8000)
- Open Records Act (KSA 21-3821)
- Public Records Act (KSA 75-3501 --- 75-3518)
- Acceptable use of the Internet (ITEC Policy 1200)
- Year 2000 Date Data Interchange (ITEC Policy 2412)
- Project Status Reporting (ITEC Policy 2500)
- Oversight of Information Technology Projects (ITEC Policy 2510)
- Project Management (ITEC Policy 2530)
- Business Contingency Planning (ITEC Policy 3200)
- Business Contingency Planning Implementation (ITEC Policy 3210)
- Technical Architecture Compliance Requirements (ITEC Policy 4010)
- Technical Architecture Change Management (ITEC Policy 4020)
- Communications Network and Systems Access Security Architecture (ITEC Policy 4210)
- Security Policy and Procedures for the KANWIN Network (ITEC Policy 4220)
- Kansas Geographic Information Systems Metadata Standard (ITEC Policy 5100)

19.8.1 Relevant KSTA Principles

- Principle #8: Data management procedures should insure that enterprise (state agency) data is maintained in a manner that provides high availability, performance, and reliability to the end user.

- Principle #9: The data management architecture must be tested and conform to generally accepted industry standards.

- Principle #11 Data management architecture should be based on commonly accepted extensible, interoperable, and scalable industry standards.

  - MDC Open Information Model 2.5, 2.6, 2.7 at: http://www.mdcinfo.com/OIM/MDCOIM11.html

  - Object Management Group standards: CORBA, UML, and Workflow at http://www.omg.org

  - Workflow Management Coalition Standards (WfMC). http://www.aiim.org/wfmc/standards/docs/tc1023v10beta.pdf

- Principle#13. Data management procedures must address and insure confidentiality of records that contain information that is limited for open/public access.

  - FERPA(students), HCFA(patients), http://www.privacyalliance.org/

- Principle #12 In all cases, agencies should carry out data management activities utilizing proven and stable IT products and processes.

19.9 Technical Product & Configuration Information

Description	Product	Notes
The Meta Data Coalition (MDC), founded in 1995, is a not-for-profit consortium of close to 50 vendors and end-users whose goal is to provide a tactical solution for metadata exchange. Participation in the MDC is encouraged and open to all vendors and end users. Information about the Meta Data Coalition is available through the MDC Web site at http://www.MDCinfo.com.		The MDC Council: • Commercial Financial Services, Inc. • ETI • MICROSOFT • NCR • PLATINUM technology, Inc. • Sybase The Technical Subcommittee currently comprises: • PricewaterhouseCoopers • NCR • ETI • MICROSOFT • Sybase • Cognos • SAS • CFS • PLATINUM technology, Inc. • Mastersoft International (MSI) • Prudential
Open GIS Consortium (OGC) Http://opengis.org		OpenGIS Specifications OGC makes the current version of the Abstract Specification public when an OGC Technical Committee Working Group issues a Report for Proposals (RFP) for engineering specifications that implement part of the Abstract Specification for particular distributed computing platforms. Development The Open GIS Consortium (OGC) has achieved consensus on several families of APIs, and some of these have now been implemented in Off-The-Shelf software.

Description	Product	Notes
The OMG was formed to create a component-based software marketplace by hastening the introduction of standardized object software. The organization's charter includes the establishment of industry guidelines and detailed object management specifications to provide a common framework for application development. Information about the Object Management Group is available through the OMG Web site at http://www.omg.org.		Founded in April 1989 by eleven companies, including 3Com Corporation, American Airlines, Canon, Inc., Data General, Hewlett-Packard, Philips Telecommunications N.V., Sun Microsystems and Unisys Corporation. In October 1989, the OMG began independent operations as a not-for-profit corporation. Through the OMG's commitment to developing technically excellent, commercially viable and vendor independent specifications for the software industry, the consortium now includes over 800 members.

Table 19.9-1: Data and Information Product Information

19.10 Futures

Future trends in Information Management Architecture are somewhat contingent on emerging technologies. Groundbreaking technology can cause major paradigm shifts that make long term planning a difficult process. The emergence of the Internet and related transports and interfaces is a prime example of the effect of breakthrough technologies. Whatever technology emerges, the following trends will continue and assume greater importance:

- Platform Independence. Relevance of a particular platform (UNIX, AS/400, Mainframe, NT) continues to diminish, due to standardized advanced programming interfaces (API). Newer software such as Java makes platform less of an issue. Operating systems and database products will handle the differences in hardware and make the server transparent to the end user.

- Universal Data Exchange. Equipment continues to evolve, however interface and data exchange standards are being developed. ISO and industry de facto standards continue to develop as needs arise.

- Faster equipment. There continue to be breakthroughs with new technology that will allow more information to be processed faster, and delivered seamlessly.

- Greater information availability. Quick delivery methods and open records laws will result in freer exchange of information with the public and other state agencies.

- Statewide Data Warehouse. A well-organized index of all appropriate state information, which is in easily understood formats, clearly documented, current, and accurate.

- More security concerns. As equipment, communications, and programming languages become more standardized, those that wish to engage in illegal or unethical practices will have more tools available. IT professional will need to be more vigilant to detect weakness in networks and prevent unauthorized access and denial of service.

19.11 Organizational & Personnel Impact

The IT community needs to lead the way in breaking down the walls that inhibit cooperation between state agencies. Greater cooperation facilitates standardization, consolidation, and increases our IT purchasing capacity to the greater benefit of all state agencies. The ultimate information goal of inter-agency cooperation is a statewide data warehouse that presents a consistent information management view to Kansas citizens, internal users, and other authorized external users.

This type of cooperation an organization comes at a price. In order for information to be universally accessible, metadata standards need to be established and adhered to. Appropriate formats for end user data need to be extracted from relational databases to facilitate ease of use and reduce complexity of multi-dimensional data system. It is clear that there will be increased demands on IT professionals to make new data systems coherent and user friendly. These evolving demands will have a profound impact:

- Greater emphasis needs to be placed on training both for our technical staff as well as our business partners in state methodology and technical skills. Having an Information Management Architecture is useless unless there is a level of understanding and expertise that takes advantage of it. State-sponsored training will present a good forum for inter-agency cooperation. New skill sets will need to be developed including metadata, new platform independent programming languages, modern databases, XML, and data warehousing. Management must come to grips with the fact that IT training

will be a continual process. Without frequent upgrades of skills, IT professionals will not be able to utilize technology to the greatest advantage for the state of Kansas.

- More IT positions will be needed. Technology has resulted in the end users of technology becoming more productive due to improvements in client interfaces, availability of information, end user software like MS Word and Excel, and e-mail. The user community is demanding more availability, better interfaces, more sophisticated programming, more products, and data cleansing. Providing these services require more staff than for previous generation users and equipment.

- More overlap and joint projects between agencies. Providing the public with functional information as opposed to agency related information would require long-term joint projects and cooperation between state agencies. The traditional way of funding certain positions may need to be reexamined as more IT professionals serve multiple agencies for extended periods of time.

20 Records Management and Preservation Architecture

20.1 Vision

When Kansas state entities conduct activities electronically, they will create and maintain reliable and authentic electronic records of those activities.

20.2 Scope and Business Rationale

20.2.1 Purpose

This chapter addresses maintaining accountability and preserving important historical records in an electronic environment. It is designed to provide guidance to users and managers of computer systems on how to:

- maintain ongoing **accessibility** of records throughout their period of retention,
- apply **retention schedules** to their electronic records,
- **manage access** to their records in a manner that ensures public access rights while also protecting confidentiality,
- address recordkeeping considerations in the **system planning and development** stage rather than waiting until the end of the records lifecycle,
- ensure the **reliability and authenticity** of records throughout their period of retention.

20.2.2 Scope

This chapter applies to public records in Kansas state and local government entities that come under the jurisdiction of the State Records Board. Policies and procedures for traditional formats of records are well-established and discussed in detail in the Kansas State Records Management Manual (2nd Edition, June 1996). This chapter applies and extends the policies and practices for records management to issues resulting from the transition from paper-based to electronic recordkeeping.

20.2.3 Definition of Records

Kansas state entities routinely create and accumulate records as they undertake government business. These records are vital to the process of managing and monitoring the use of state resources, and they provide a historical record of decisions, changes, and outcomes. Records have a significant role in the democratic process in that they:

- provide evidence to support the rule of law,
- support the accountability of government administration,
- are evidence of the interactions between the people of Kansas and their government, and
- have value in documenting the history and culture of Kansas.

The Government Records Preservation Act defines records in terms of their function and their relationship to the transaction of official business. According to K.S.A. 45-402(d) (emphasis added):

> Government record means all volumes, documents, reports, maps, drawings, charts, indexes, plans, memoranda, sound recordings, microfilms, photographic records and other **data, information or documentary material, regardless of physical form or characteristics, storage media or condition of use, made or received by an agency in pursuance of law or in connection with the transaction of official business or bearing on the official activities and functions** of any governmental agency. Published material acquired and preserved solely for reference purposes, and stocks of publications, blank forms and duplicated documents are not included within the definition of government records.

Records can be created and stored using many different media and formats, including paper-based files or computer systems, on a single medium or as multimedia. Records can also be transferred from one medium to another and from one context to another through copying, imaging or digital transfer.

Not all data in electronic information systems constitute records. Records have a distinct legal and administrative status that not all information and documents have. They must be managed as important resources with special requirements that may be distinct from other information resources. Electronic records management principles are relevant whenever computer systems are used not only to process information but also to provide reliable and authentic evidence that given activities or transactions have occurred.

20.3 Context & Diagrams

20.3.1 Best Times to Address Electronic Records Management

It is in the best interest of state entities to address electronic records management issues as soon possible. Since effective management of electronic records depends so heavily on the information systems involved, however, one will have the most options for managing electronic records effectively by identifying recordkeeping requirements when new systems are designed or when existing systems are upgraded.

Business Process Redesign

Process analysis and redesign often identify problems which could be alleviated through new workflow procedures and/or information systems, e.g. areas where electronic records are printed and filed unnecessarily because there were no provisions in the system to capture records electronically and transfer them to an electronic recordkeeeping system.

20.3.1.1 System Design and Procurement

Several aspects of recordkeeping should be considered during the system design and procurement process. Does the state entity require the system to support electronic recordkeeping, or does it plan to produce and file in hard copy all of the records that the system generates? If the system is expected to support electronic recordkeeping, then some customization of commonly available software may be needed.

Special measures may be needed for routing documents from the active information processing environment to a recordkeeping system where records can be stored but not altered after they have been filed electronically. Processes need to identify the official copy and handle version control. If the retention requirements are identified when the system is designed, routines can be designed for automatic purging of obsolete documents. If the system will store records with enduring value, a method will be needed for migration or export of the records to the next generation of technology.

20.3.1.2 Replacement and Upgrading of Information Systems

Analysts can review the recordkeeping aspects of the system that is being phased out and use that analysis to identify opportunities for improvement. If users had difficulty identifying and retrieving the most current version of a document in the old system, for example, some form of version control may be needed in the new system. If users were reluctant to rely on the electronic records and instead printed and filed large volumes of paper records, the new system could incorporate better organization of records and better retrieval capabilities. If the old system was cluttered with obsolete files, the new system

could be designed to automatically delete or transfer to offline storage specific types of files after a given time period. If users were not willing to trust the electronic versions of records, more effective authentication and system security measures could be implemented.

One important consideration when systems are replaced or upgraded is whether any of the electronic records stored in the old system need to be retained and migrated into the new system. This process can be routine if the records are stored in a simple structure or in a format that is compatible with the new system and if they are readily identifiable and well described, but detailed analysis may be necessary to identify which records need to be retained and to determine how to transfer them to the new system.

20.3.2 Creating Electronic Records

Kansas state entities create records in order to:

- produce evidence of individual and corporate performance,
- account for the use of public resources,
- document decision making processes in accordance with the law,
- comply with statutes, regulations, instructions, guidelines and other rules that require state entities to create records,
- preserve the corporate memory of the state enterprise and track business transactions over time,
- enable the government to protect its interests and to substantiate the rights and entitlements of individual citizens,
- ensure that records of significant government policies and activities are kept for posterity, and
- provide a record of communications within and between state entities and between the government and its citizens.

It is important that state entities determine how and why electronic records are being created. Many of the considerations laid out in this chapter - capture of appropriate content, creation of metadata, declaration of record type - are best addressed at the point of record creation or very shortly thereafter.

20.3.3 Capturing Electronic Records

Strategies for capturing electronic records will differ, depending on the opportunities presented by a state entity's hardware and software environment. Locations at which records can be captured include software layers (especially suited to open systems environments) and at every interface

between hardware components through which the relevant data passes. The technological environment will influence the decisions as to whether records are captured through:

- the user interface layer,
- modification of the application software,
- the operating system,
- the application program interface (API), or
- the front end to a corporate filing system.

The organizational environment will also influence the point at which records are captured. This will include perceptions about what constitutes a record, assignment of responsibility, state entity requirements to create records, and staff understanding of the technology involved.

Regardless of the approach a state entity takes, it must be able to identify specific information objects (e.g. documents, email messages, database entries) as records and somehow distinguish between the types of records to which different business and retention requirements must be applied. Possible approaches include:

- Business transaction information is identified in an "envelope" or file header, so the file does not need to be opened to be identified.
- The record creator is responsible for capturing his or her own records and assigning management practices to them at the point of creation. This could be implemented as a screen the user fills in before documents can be saved or messages can be sent.
- A user interface is designed so that users can choose between a number of icons representing business tasks or style templates, e.g., "send policy" or "make appointments." The choice of icon can engage the appropriate application, distribution lists, style sheets and records disposal authorities. The sender thus affects scheduling but need not make conscious decisions about assigning retention periods to records.

20.3.4 Identifying Electronic Records

State entities traditionally have used records surveys and inventories to identify which records they maintain and to decide what to do with those records. In an electronic context, surveys of physical storage media (e.g. tape libraries or workstation hard drives) do not provide much useful information for determining which records exist or for deciding what to do with them. As explained in the Data Management Architecture (Chapter 18), in order to enhance performance and convenience, most information systems make use

of redundant data. Instead of attempting to inventory all of this data that exists at any one time, electronic records management requires the identification of state entity functions, processes, transactions and activities to be documented. Once these have been identified, it will be possible to determine which data and associated metadata must be retained to serve as an official record.

20.3.5 Managing Electronic Records

State entities need ready access to the right information at the right time to provide services and make informed decisions. An important part of that process is gathering information together to form the basis for decision making. Another part of the process is internal and external communication using various technologies. This communication process invariably involves conducting some form of business transaction (development of policy, delivery of benefit, ordering or paying for a product or service) which needs to be documented. The means by which state entities choose to conduct these business transactions invariably involve oral, written and/or electronic communication methods. In all cases, the objective is to conduct the business transaction satisfactorily and to maintain a record of what transpired for future reference.

When conducting transactions electronically, the first challenge is to maintain records in a way which will enable retrieval of all documents relevant to a transaction when they are needed. The second challenge is to ensure that the records are not retained for any longer than necessary, in order to avoid both overloading systems and to avoid indiscriminate dumping. A special problem with electronic records is that they lack familiar physical and visual clues about their origins, such as official letterhead, or their authenticity, such as written signatures. Special measures must be taken to ensure that they are also reliable and authentic.

Paper recordkeeping systems traditionally have been employed to file letters, minutes, reports, spreadsheets, invoices, notes, etc. These systems employ classified and indexed files at a subject or transaction level to consolidate and co-locate the documents generated or received in the course of a business activity. Separate folders provide a business context and link the individual documents to a particular transaction and into the wider state entity recordkeeping system. In recent years, state entities have adopted records management, document management, workflow and imaging software. Regardless of the technology, however, the objective remains the same: capture records so that they can be easily retrieved at a later date, understood, and interpreted as evidence of what transpired in a state entity.

By reducing records to their essential characteristics, we can allow for the existence of records, regardless of the current technology. Systems must link the content of a record to its administrative or business context. In electronic environments, the essential characteristics of records rarely sit neatly together in a single, format-based package. Though all of the elements of a virtual

record may exist within a single computer file, they may also be distributed across the entire state network. The integrity of these elements and the links between them are much more important than where they physical reside. If one is not able to place records in their appropriate administrative context, then they have seriously diminished value as evidence.

20.3.5.1 Full and Accurate Records

Records should be full and accurate to the extent necessary to:

- facilitate action by current and future employees, at all levels;
- allow for proper scrutiny of the conduct of business by anyone authorized to undertake such scrutiny; and
- protect the financial, legal and other rights of the state, its clients and anyone else affected by its actions and decisions.

20.3.5.2 Essential Characteristics of Records

Full and accurate records must posses the following three essential characteristics:

- **Content** -- that which conveys information (e.g. text, data, symbols, numerals, images, and sound).
- **Structure** - appearance and arrangement of the content (e.g. relationships between fields, entities, language, style, fonts, page and paragraph breaks, links and other editorial devices).

Context - background information that enhances understanding of technical and business environments to which the records relate (e.g. metadata, application software, logical business models) and the origin (e.g. address, title, link to function or activity, state entity, program or section).

In order for records to serve as evidence, these three essential characteristics must be maintained. Whenever one of the characteristics is altered, the ability of records to accurately reflect the activities of a state entity is diminished. This means that records must:

- have information content that is (and continues to be) an accurate reflection of what actually occurred at a particular time in the function, activity or transaction in question;
- be able to be reconstructed electronically when required, so that each component is brought together as a whole and presented in an intelligible way;

- be able to be placed in context so that the circumstances of its creation and subsequent use by a state entity or person can be understood in conjunction with its information content; and

- have been officially incorporated (either actively or passively) into a state entity's or person's recordkeeping system.

A major difference between electronic records and those on traditional media is that electronic records are not human-readable. Their physical appearance alone is not sufficient to determine their origin, purpose, uses or other aspects of the context in which they were created and maintained. Maintaining content, structure and context of electronic records is, therefore, both more vital and difficult than with traditional records. Meeting these conditions requires high quality records management and a sustained commitment.

20.3.6 Recordkeeping Systems Defined

Recordkeeping systems are those systems that capture, manage and provide access to records over time. Records are often accessed solely for their informational content, in which case they function the same as any other document or information source. Records are kept, however, to provide evidence of functions, activities and transactions, i.e., the business process. Recordkeeping systems are different from generic information systems in that they maintain linkages to the activities they document and preserve the content, structure and context of the records.

Unlike most other computer information systems, recordkeeping systems often must accommodate records that exist in more than one format (e.g. parallel paper case files and electronic case management systems). Recordkeeping systems should be able to identify all records, active and inactive, and the version of the computer software that supports access. They should be able to identify records stored off-line and off-site and on all media.

20.3.7 Building the Essential Characteristics into Recordkeeping Systems

The realities of modern administrative practice often can be impediments to effective recordkeeping. The pressure of the moment and the thought that documentation can wait have increasingly become a standard feature of modern organizations. The introduction of a greater commercial and service orientation in the public sector has created a culture which is focused on outcomes, sometimes to the detriment of documentation.

Effective electronic records management is not a goal to be attained at the expense of state or state entity outcomes but is instead a necessary component of those outcomes. When successful outcomes are well documented, they can be sustained over time, accurately reported to the citizens of Kansas, and potentially reapplied across the state enterprise. When outcomes are not well documented, however, the state of Kansas can neither leverage its past successes nor avoid repeating its past failures.

The systematic creation and keeping of records have been undermined by the move away from centralized filing systems, the introduction of risk management, outsourcing, and the increasing use of technology in the administrative process. This is not to suggest the state return to the centralized and resource-intensive practices of the past. Rather, state entities should put systems in place which meet their accountability requirements without detracting from the benefits provided by modern technology and organizational change. When the system will support or provide services for several state entities, those entities involved should work together to ensure that all of their respective recordkeeping requirements will be met.

The longer records are maintained, the more difficult it becomes to fully maintain their content, structure and context. In the process of upgrading, converting or migrating data to accommodate new systems, one or all of the essential characteristics of records may be compromised in some way. If the practices recommended in this chapter are applied to the design, implementation and management of information systems, however, this loss of essential characteristics can be minimized and state entities can make better decisions about which characteristics warrant the resource commitment to maintain.

20.3.7.1 The Importance of Open Standards

Data management, interchange, interoperability, migration and ongoing accessibility all depend on the adoption of open standards. Although some components of state computer information systems inevitably will be proprietary, electronic records management should not be dependent upon the software or hardware of one particular vendor. Whenever feasible, file formats, protocols and other system specifications adopted by state entities should be those developed and adopted by recognized standards bodies. Since the requirements for fulfilling these standards are both publicly documented and generally supported by more than one vendor, state entities that adopt them will be much less likely to find themselves stuck with valuable but inaccessible records than will those that adopt more closed systems. The appropriate standards body will depend upon the nature of the technology involved, but three particularly important sources of standards relevant to electronic records management are the International Organization for Standardization (ISO), Internet Engineering Task Force (IETF) and World Wide Web Consortium (W3C).

20.3.7.2 Content

In order to maintain record content, state entities should follow best practices in the information technology profession for data integrity. Systems should be in place to ensure that:

- the identity of a record's creator is verified (through the use of a password and possibly encryption),
- permission to read and write files is appropriately restricted,
- periodic system audits are conducted,
- data transmission includes data error checking and correction,
- data are regularly backed up, and
- data on offline media such as magnetic tape are regularly refreshed to avoid catastrophic loss of data due to medium degradation.

Data also should be encoded in such a way that the bits will continue to be readable over time. Records that contain American Standard Code for Information Interchange (ASCII) text provide an easy migration path with respect to content as long as ASCII remains an accepted base standard. Open Systems Interconnection (OSI) standards for other forms of content, e.g. Tag Image File Format (TIFF) for images, also should be considered for long-term retention of records. For nontextual materials, it is important to distinguish between record copies and convenience copies. If a paper document has been digitized, for example, a state entity may store a master copy of the document as a high-resolution TIFF image for preservation purposes but provide online access to a lower resolution Joint Photographic Experts Group (JPEG) or Graphics Interchange Format (GIF) image that serves only as convenience copy for easy reference.

As previously stated, the management of records should not be restricted to records that reside only on certain media types. The records of business processes may span different media and multiple systems. Business decisions to restrict record creation to certain media should be clearly articulated and communicated to staff. Recordkeeping systems should be designed to enable access to the complete record without hindrance. Where multiple recordkeeping systems are in place, links should be provided for records that span these multiple systems.

20.3.7.3 Structure

Recordkeeping systems need to capture and maintain information about the structure of records either as an integral part of the metadata associated with the records or in separate formal documentation. In many ways, structure is more difficult to maintain than content and is often neglected.

The simpler the record structure, the easier it is to preserve the record over time. As with the other characteristics of records, it is also best for record structure to be based on open standards. See the Data Structure entry under 20.7 for examples of open standards for document structure.

20.3.7.4 Context

If the content of a record becomes separated from key information about the entity or entities who made it, the time, place and reasons for its creation, and its relationship to other records, its value as a record is severely diminished or lost. Its contents may still be of interest, but the record will have no value as evidence unless it can be placed in context. Contextual information, therefore, is information about the records and the administrative environment in which they were created and maintained. It can range from high-level information such as the name and location of the entity or entities that created the record to more detailed information such as the date the record was made.

The depth of contextual information required will vary depending on the expected users and their level of knowledge. In the case of permanent records, more details will be necessary to enable future audiences to make sense of the records and place them in context. What is commonly known and assumed by today's records creators may not be readily evident to future users.

Recordkeeping systems need to maintain and provide access to information about the business and administrative context in which records were created and used. For computer systems developed by information technology professionals, system design documentation, data dictionaries and related business documentation are fundamental to providing context for records that are held in those systems. Active data dictionaries - lists of all files in a database management system, the number of records in each file, and the names and types of each field - and computer-aided software engineering (CASE) tools - software that provides a common development environment for programming teams - automate much of the process of keeping metadata authentic.

Maintaining the context of records created and managed outside of systems developed by information technology professionals is more difficult. The ubiquity of personal computers allows records to be created, modified, copied, transmitted and deleted, often with little regard for business and legal records management requirements. Even if records are managed appropriately on an individual workstation, their existence may not be known to other users, and the contextual information may be inadequate for future retrieval.

Consideration needs to be given to assigning and preserving meaningful document names, author, work group and organizational identifiers, designating whether records are draft or final versions and linking them to other documents or information objects. Off-the-shelf software exists to address these problems. Alternatively, if records cannot be supported in an electronic environment, they will need to be output to a recordkeeping system based on paper, microfilm or some other analog medium.

Contextual Information Provided by State Entities

Contextual information needs to be collected, structured, and maintained from the time records are created. This involves identifying and labeling (or tagging) records and linking them to contextual information (i.e., keeping records about records). In some cases this can be achieved by embedding key contextual information into the metadata or electronic records themselves. The more that electronic records can be made self-describing the less need there is for maintaining separate information.

State entities can use some combination of the following methods to incorporate records management activities into their information systems:

- Purchase and implement specific records management software (see examples in 20.9 below).
- Configure existing software to include records management functions.
- User-based management. The users of information systems can manually engage in electronic records management functions.

Regardless of which of the above methods state entities adopt, the Archives encourages them to maintain contextual information relating to the:

- entity or entities that recorded or maintained the records,
- other entities that are, or have been, associated with the records,
- purpose of the records in fulfilling state functions;
- age of the records,
- time period to which the records relate,
- frequency with which the records are, or will be, used,
- value or significance of the records in relation to the functions of the state entity,
- recordkeeping system used in relation to the records,

- relationship (if any) between the records and other records or materials, and

- existence of any law, agreement, practice, procedure, arrangement or understanding affecting the records.

Such contextual information, while desirable for all records, is especially important for higher value records. While such contextual information is absolutely necessary for long-term retention of electronic records, it can also improve the quality of records in active use, support information sharing, and enhance their quality as evidence.

Contextual Information for Inter-Entity Transfer

When electronic records are transferred from one state entity to another following changes in government administrative arrangements or are transferred to the Archives, it is essential that they are transferred with sufficient metadata and contextual information. State entities that take on the care and preservation of electronic records under such circumstances need to insist that the relinquishing entity supply adequate contextual information, system documentation and metadata at the time of transfer. Because of the risks involved, state entities transferring electronic records between themselves, either directly or through a contracted service provider, should follow verification procedures. This process is increasingly happening in real time. Systems for interchange must ensure not only the transfer of data but also sufficient metadata.

Contextual Information Gathered by the Archives

To manage records and determine their appropriate retention and disposition, information will be gathered about records and maintained in a database by the Kansas State Historical Society (KSHS). Records will be classified according to their record series, which is a group of records normally used or filed as a unit that relate to a particular subject or result from the same activity.

For an individual record or a series of related records the KSHS gathers information about:

state entity

- title of the state entity (or entities) which created the records,

- dates between which the entity operated,

- purpose of the entity and the functions and legislation it administered at the time,

- location of the entity, and

- outline of the state entity's development, history, internal structure and relationship to other entities.

Series

- title of series (e.g., business process or function) to which the record belongs,
- date range of the series
- content and purpose of the series,
- system of arrangement or control of the series (e.g., retrieval system, indexing system),
- quantity of records in the series,
- previous and subsequent series (if any) that document the same or similar functions,
- controlling, controlled and related series (e.g., indexes, data dictionaries),
- identity of other state entities or persons that have had custody of the records,
- relevant disposition schedules and actions taken, and
- statutes, regulations, and policies governing access.

Archives staff obtain this contextual information from a range of sources, including records disposition schedules, transfer documentation, direct physical examination of record items, and research through published and state entity sources. The vulnerability of electronic records is such that state entity staff responsible for their creation and management must now take an active role in ensuring that sufficient contextual information is gathered so it can be provided to the Kansas State Archives in the event of a transfer of permanent records to the Archives.

20.3.8 The Problem of Legacy Records

The Section 3.10 of the Kansas State Technical Architecture (KSTA) divides the lifecycle of information technology into four phases: introduction/emerging, growth/acceptance, stability and twilight. The KSTA recommends the adoption of products that are currently in either the second or third phase of the technology lifecycle. Regardless of what phase a system is in at the time of implementation, however, it will eventually enter the twilight phase. In order to maintain access to the records on these older systems, state entities must take measures to either continuously support those systems or migrate the records to newer systems. The record lifecycle is thus tightly connected to the technology lifecycle. In short, electronic records live and die with the systems that support them.

This dependency becomes a major problem in the case of legacy records, which are records that rely on legacy systems. Legacy systems are those systems that were designed using hardware and software systems that are rapidly becoming obsolete or are no longer supported by their vendors. Legacy systems are a significant problem for organizations that rely on older, proprietary systems and technology because it is difficult to migrate either the functionality or the data to new generations of systems.

From a records management and archival perspective, legacy systems create problems when they are being used to store and retrieve records that need to be kept beyond the useful life of the system itself. There are a variety of methods that can be used to extract records from legacy systems, ranging from simply printing records to paper or microforms to using sophisticated extraction tools. Because migration is expensive, regardless of the approach used, it is important to thoroughly analyze the records and their retention requirements so that only those records that are needed for future use or required to be kept by law are migrated.

The most effective way to address the long-term retention of electronic records is to ensure that they never become legacy records. If state entities follow the recommendations in these guidelines about the capture of system metadata and thorough documentation of information systems, then electronic records will be much easier and cheaper to maintain over time. Of course, metadata that identifies the system requirements for accessing electronic records will be of no use if future users do not also have the tools needed to satisfy those requirements. This is why state entities should adopt open standards whenever possible. This will increase the chances that records can survive the transition to a new system without the need to significantly alter them in the process.

Even if state entities adopt open standards, however, cases will arise in which state entities no longer have access to software or hardware that can support a given standard or set of standards. In these cases, a factor that can greatly facilitate support for and/or migration from twilight systems is access to their source code, the sequence of statements that are written by and understandable to a human programmer. Without access to source code, state entities are more dependent on software vendors -- who may go out of business or require the purchase of a prohibitively expensive new release of their product -- to maintain the means to access their electronic records. Having access to the source code allows the entity using the software to contribute to its further development and more easily develop other software that interacts with it. There are several ways that state entities can ensure access to source code:

- Develop software internally, then maintain and document the source code.

- Make use of open-source software (OSS). OSS is software for which the source code is freely and publicly available, though the specific licensing agreements vary as to what one is allowed to do with that code. When using OSS, it is important to ensure that the software has been sufficiently documented by its developers.

- Specify in contracts with vendors that they must provide source code along with the binary code of their software and any upgrades. Restrictions may be placed on how the entity can manipulate, reuse or distribute the source code.

- Make arrangements with a trusted third party to hold the source code in escrow. There are a number of companies that provide such services, and escrow agreements can specify that access to the source code only be allowed under specific conditions.

20.3.9 Stewardship of authentic electronic records - evidence

State entities use diverse systems and technologies to create, maintain and reproduce records. Increasingly, many of those documents are created and maintained in electronic form. While information technologies enable state agencies to streamline recordkeeping practices and reduce records creation and storage costs, they also present new challenges to establishing the authenticity and the admissibility of records. Information systems and records management policies must ensure that agencies produce and maintain full and accurate records that are acceptable for legal, audit, and other purposes.

Managing and maintaining authentic electronic records in a complex, changing technical environment is a challenging undertaking that requires cooperation and coordination within and, increasingly, among state entities. A state entity's business managers, records staff, legal counsel, and information technology personnel all must be involved in ensuring the legal authenticity of records. Advice also may be requested from the Kansas attorney general and State Archives.

Evidence that is introduced in legal proceedings is subject to case law, Kansas and federal rules of evidence. Courts traditionally are prepared to rule on the admissibility of records created by common information processing methods and technologies, such as writing, typing, photocopying and microfilming. However, records produced or reproduced using newer technologies, such as digital imaging, workflow and document management systems, groupware, electronic data interchange (EDI), and electronic

commerce are uncharted waters for the jurist, because recognized standards for the implementation and use of these technologies are not yet settled in the legal practice. State entities must exercise extreme care when implementing electronic systems to ensure that these systems are reliable and that they produce records that will be legally sufficient.

Courts are more likely to admit electronic records as evidence if agencies have taken the following precautions in the design and management of their recordkeeping systems:

- Use the recordkeeping system consistently and in the normal course of business,
- Develop and follow written policies and procedures,
- Provide training and support,
- Develop an adequate system of controls,
- Conduct routine tests of system performance,
- Test and document the reliability of hardware and software,
- Provide adequate security,
- Establish controls for accuracy and timeliness of input and output, and
- Create, maintain, and retain comprehensive system documentation.

Any of the measures recommended for good systems design, system maintenance, and electronic recordkeeping also enhance the quality of electronic records as evidence.

Finally, evidence, as a concept, is not confined to legal contexts. Within business and public sector environments, the evidence from previous actions and decisions is used as precedent for the formulation of new decisions and actions. Organizations keep records as evidence or proof that an activity or transaction did or did not occur. Beyond this more immediate use, researchers also use records as historical evidence on which to base their conclusions.

20.4 Principles

- State entities should maintain ongoing accessibility of records throughout their period of retention.
- State entities should take measures to ensure the accurate and consistent application of retention schedules to their electronic records.

- Electronic recordkeeping systems should be based on open standards, whenever practical.

- State entities must manage access to their records in a manner that ensures public access rights while also protecting confidentiality.

- Recordkeeping considerations should be addressed in the system planning and development stage rather than waiting until the end of the records lifecycle.

- State entities should take measures to ensure the reliability and authenticity of records throughout their period of retention.

20.5 Goals

- Kansas state government will maintain ongoing accessibility of records throughout their period of retention.

- Ensure public access rights while also protecting confidentiality.

- State and state entity systems, policies and procedures will reflect and address recordkeeping requirements.

20.6 Best Practices & Processes

- Refresh physical storage media periodically to compensate for media degradation.

- Either store the supporting application software with data or convert it to new formats as systems change.

- Migrate data to new systems (desktop operating systems, network operating systems, enterprise management systems, etc) as they are implemented.

- Maintain metadata and documentation to identify appropriate retention periods.

- Use some combination of records management applications, user-based management, and extensions to existing applications and operating systems to both associate and apply retention schedules with the appropriate records.

- Continue to monitor and participate in the Kansas State Technical Architecture process, in order to ensure the adoption and implementation of appropriate standards.

- Capture and maintain system metadata for records that specifies appropriate access permissions.

- Provide online access to both active and inactive records, when appropriate.

- Maintain active communication between those responsible for electronic recordkeeping and those responsible for satisfying Open Records Act requests.

- When feasible, implement automatic measures for redacting confidential data from otherwise public records, rather than printing out documents and then redacting manually.

- Create and maintain metadata that adequately reflects the content, context and structure of records as they were originally created.

- Create and maintain system documentation.

- Maintain accurate system logs.

- Use authentication to identify the users of the system.

- Include recordkeeping requirements in project plans and Requests for Proposals (RFPs) for new projects.

- Develop retention and disposition schedules, and recordkeeping plans when appropriate, as part of the system development process.

- Include records capture, identification, management and retention scheduling in the business rules of new systems.

- Restrict write permissions on official records.

- Use some combination of records management applications, user-based management or extensions to existing applications and operating systems to create sufficient structural and contextual metadata at the point of record creation or shortly thereafter.

- When data is created and maintained in the course of a state activity that requires documentation, capture and maintain it as a record unit.

- Either encapsulate metadata into information objects themselves or provide appropriate links between information objects.

- Maintain appropriate links between database fields.

20.7 General Standards

Category	Emerging	Current	Twilight
Accessibility	• Authoring Tool Accessibility Guidelines • User Agent Accessibility Guidelines	• Web Content Accessibility Guidelines	
Application Program Interfaces (API)	• Document Object Model (DOM) • Java Database Connectivity (JDBC) • Open Document Management Association API (ODMA) • Simple API for XML (SAX)	• Open Database Connectivity (ODBC)	
Archival Description	• General International Standard Archival Description (ISAD(G))	• Endoded Archival Description (EAD)	
Character Encoding	• Unicode	• American Standard Code for Information Interchange (ASCII) • ISO Latin-1 (ISO-8859-1)	• Extended Binary-Coded Decimal Interchange Code (EBCDIC)
Data Content		• Archival Moving Images: A Cataloging Manual • Archives, Personal Papers and Manuscripts (APPM) • Anglo-American Cataloguing Rules, 2nd ed. (AACR2) • Content Standard for Digital Geospatial Metadata (FGDC-STD-001-1998) • Draft Interim Guidelines for Cataloging Electronic	

Category	Emerging	Current	Twilight
		Resources • Graphic Materials: Rules for Describing Original Items and Historic Collections • Oral History Cataloging Manual • Subject Cataloging Manual: Subject Headings, 5th ed.	
Data Interchange	• Internet Protocol, Version 6 (IPv6) • Multiprotocol Label Switching (MPLS) • Protocol Independent Multicast-Sparse Mode (PIM-SM) • Wireless Application Protocol (WAP)	• Asynchronous Transfer Mode (ATM) • File Transfer Protocol (FTP) • HyperText Transfer Protocol (HTTP) • Internet Protocol, Version 4 (IPv4) • Multipurpose Internet Mail Extensions (MIME) • Point-to-Point Protocol (PPP) • Post Office Protocol, Version 3 (POP3) • Transmission Control Protocol (TCP)	• AppleTalk • DECnet • Synchronous Data Link Control (SDLC) • Systems Network Architecture (SNA) • x.25
Data Semantics	• Dublin Core • Namespaces in XML • Resource Description Framework (RDF) Model and Syntax • Resource Description Framework (RDF) Schemas		

Category	Emerging	Current	Twilight
	• Semantic Annotations in HTML • Topic Maps (ISO 13250) • XML Schema Part 1: Structures • XML Schema Part 2: Datatypes		
Data Structure	• Data Documentation Initiative (DDI) • DocBook • Global Information Locator Service (GILS) • HL7 Reference Information Model • MAchine Readable Cataloging (MARC) • Mathematical Markup Language (MathML), Version 1.01 • Structure for the identification of organizations and organization parts (ISO 6523) • Text Encoding and Interchange (TEI P3) • Voice Markup Language (VoxML) • Wireless Markup Language (WML) • eXtensible HyperText Markup Language (XHTML)	• HyperText Markup Language (HTML), Verion 4.01	
Data Syntax	• Canonical XML, Version 1.0 • eXtensible Markup	• Comma-Separated Value (CSV)	• Data Interchange

Category	Emerging	Current	Twilight
	Language (XML) • XML Fragment Interchange • XML Information Set	• Directory Interchange Format (DIF) Formal Syntax Specification v7.0 • Standard Generalized Markup Language (SGML)	Format (DIF)
Data Values	• AGIFT (Australian Government's Interactive Thesaurus) • Keyword AAA • United Nations Educational, Scientific and Cultural Organization (UNESCO) Thesaurus	• Art and Architecture Thesaurus (AAT) • Dates and Times (ISO 8601) • Dictionary of Occupational Titles • Global Legal Information Network (GLIN) Subject Headings • ISAAR - CPF (Corporate Bodies Persons and Families) • Legislative Indexing Vocabulary (LIV) • Library of Congress Name Authority File (NAF) • Library of Congress Subject Headings (LCSH) • Medical Subject Headings (MeSH) • National Aeronautics and Space Administration (NASA) Thesaurus • Resource Description Framework (RDF) • Revised Nomenclature for Museum Cataloging	

Category	Emerging	Current	Twilight
		• Roget's Thesaurus • Rules for the Construction of Personal, Place and Corporate Names • Specification and standardization of data elements (ISO 11179) • Thesaurus for Graphic Matierals (TGM): Subject Terms (TGM1) and Genre and Physical Characteristic Terms (TGM2) • Thesaurus of Geographic Names (TGN) • Union List of Artists Names (ULAN)	
Data Transformation	• XSL Transformations (XSLT), Version 1.0		
Digital Signatures and Authentication	• Certificate Issuing and Management Components Protection Profile • Federal Public Key Infrastructure (FPKI) • Public-Key Infrastructure (X.509) (PKIX) • PKI Practices and Policy Framework (X9.79) • XML-Signature Requirements • XML-Signature Core Syntax and Processing	• Kerberos • ITU-T X.509 (ISO/IEC 9594-8), Version 3	
Encryption	• Advanced Encryption Standard (AES)	• Data Encryption Standard (DES) - (ANSI X3.92,	

Category	Emerging	Current	Twilight
		X3.106 and FIPS 46, 81) • SHA-1 (FIPS180-1, ANSI 930-2, ISO/EC 10118-3)	
File Naming and Hierarchy		• Filesystem Hierarchy Standard (FHS) • Universal Naming Convention (UNC)	
Geographic Information Systems (GIS)	• OpenGIS Abstract Specification		
Images	• Wireless BitMap (WBMP) • Computer Graphics Metafile (CGM) • Initial Graphics Exchange Specification (IGES) • International Color Consortium (ICC) Specification • JPEG 2000 • JPEG Network Graphics (JNG) • Portable Network Graphics (PNG) • Multiple-image Network Graphics (MNG) • Scalable Vector Graphics (SVG) • WebCGM	• Graphics Interchange Format (GIF) • Joint Photographic Experts Group (JPEG) • Tag Image File Format (TIFF)	
Information Retrieval	• Australian Government Locator Service (AGLS) • Common Indexing Protocol (CIP)	• Common Gateway Interface (CGI) • Structured Query Language (SQL)	

Category	Emerging	Current	Twilight
	• Global Information Locator Service (GILS) Application Protocol • XML Matching and Structuring language (XMAS) • XML Query Language (XQL) • XML Query Requirements • Information Retrieval: Application Service Definition and Protocol Specification (Z39.50)		
Metadata for Recordkeeping and Preservation	• Recordkeeping Metadata Schema (RKMS) - Strategic Partnership with Industry, Research and Training (SPIRT) • Recordkeeping Metadata Standard for Commonwealth Agencies - Australia • Victorian Electronic Records Strategy (VERS) Metadata Specification	• Functional Requirements for Evidence in Recordkeeping • Metadata For Digital Preservation - Consortium of University Research Libraries Exemplars in Digital Archives (CEDARS) • Reference Model for an Open Archival Information System (OAIS) - Consultative Committee for Space Data Systems (CCSDS)	
Multimedia	• Advanced Authoring Format (AAF) • Advanced Streaming Format (ASF) • MPEG Audio Layer 3 (MP3) MPEG 7 Synchronized Multimedia Integration Language (SMIL) Virtual Reality Modeling Language (VRML) Visual XML (VXML)	• Moving Picture Experts Group (MPEG-1, MPEG-2 and MPEG-4) • Musical Instrument Digital Interface (MIDI)	

Category	Emerging	Current	Twilight
Object Modeling and Interchange	• Common Object Request Broker Architecture (CORBA) • Document Object Model (DOM) • Internet Inter-ORB Protocol (IIOP) • Unified Modeling Language (UML)		
Rating and Filtering	• A P3P Preference Exchange Language (APPEL) • Platform for Internet Content Selection (PICS) • Platform for Privacy Preferences (P3P), Version 1.0		
Records Management Software	• Design Criteria Standard for Electronic Records Management Software Applications (DoD 5015.2-STD)		
Records Management Strategies	• Australian Standard for Records Management (AS 4390)		
Resource Identifiers and Links	• Digital Object Identifier (DOI) • Directory Services Markup Language (DSML) • Common Name Resolution Protocol • Handle System • XML Pointer Language (XPointer) • XML Path Language (XPath)	• Domain Name System (DNS) • Uniform Resource Locator (URL) • Lightweight Directory Access Protocol (LDAP) • x.500	

Category	Emerging	Current	Twilight
	• XML Linking Language (XLink) • Uniform Resource Name (URN) • Persistent Uniform Resource Locator (PURL)		
Scripting	• ECMAScript	• JavaScript • JScript • Perl	
Secure Transfer	• Encryption using KEA and SKIPJACK • IP Security (IPsec) • KeyNote Trust-Management System Version 2 • RSVP Operation Over IP Tunnels • Secure Electronic Transaction (SET) • Secure Multipurpose Internet Mail Extensions (S/MIME) • Simple Key management for Internet Protocols (SKIP)	• SSH Protocols and Secure Shell • Secure Sockets Layer (SSL)	• Secure Hypertext Transfer Protocol (S-HTTP)
Style Sheets	• eXtensible Stylesheet Language (XSL)	• Cascading Style Sheets (CSS)	• Document Style Semantics And Specification Language (DSSSL)
Wrappers & Mediators	• Mediation of Information Using XML (MIX)		

Category	Emerging	Current	Twilight
	• Warwick Framework • Universal Preservation Format (UPF)		

20.8 Related Policies & Procedures

20.8.1 Statutes

- Americans with Disabilities Act (ADA) - (42 USC 12101, 28 CFR 35.160)
- Child Online Privacy Protection Act (COPPA)
- Computer crime; computer password disclosure; computer trespass (KSA 21-3755)
- Computer Security Act of 1987 (40 USC 759, Public Law 100-235)
- Digital Milenium Copyright Act (DMCA) - (Public Law 105-304)
- Electronic and Information Technology (Title IV, Section 508)
- Electronic Communications Privacy Act (18 USC 2701)
- Electronic Freedom of Information Act (E-FOIA) - (amdendment to 5 USC 552)
- Freedom of Information Act (FOIA) - (5 USC 552)
- Government Paperwork Elimination Act (GPEA) - (Title XVII)
- Government Records Preservation Act (KSA 45-401 through KSA 45-413)
- Kansas Acts Against Discrimination (KSA 44-1001 et seq)
- Open Records Act (KSA 45-215 through 45-223)
- Public Records Act (KSA 75-3501 through 75-3518)
- Records made on Electronically-accessed Media; Authorization, Conditions and Procedures, Application, Notice to State Records Board (KSA 45-501)

- Standards for Electronic and Information Technology (Notice of Proposed Rulemaking on Standards for Electronic and Information Technology implementing Section 508 of the Rehabilitation Act)
- Tampering with a Public Record (KSA 21-3821)
- Telecommunications services of certain state agencies; extension to certain private, nonprofit agencies or governmental entities; records of services (KSA 75-4709)
- U.S. Copyright Act (17 USC 101 - 810)

20.8.2 Regulations

- Electronic Records Management - Title 36, Code of Federal Regulations (CFR), Chapter XII, Part 1234
- Federal Rules of Evidence
- General Records Retention and Disposition Schedule for State Agencies (KAR 53-3-1)
- General Retention Schedule (GRS) 20 - National Archives and Records Administration
- Records Maintained On Individuals - Privacy Act (1999 CFR Title 10, Volume 4)
- Records Officer (KAR 53-4-1)

20.8.3 Policies

- Acceptable use of the Internet (ITEC Policy 1200)
- Business Contingency Planning (ITEC Policy 3200)
- Business Contingency Planning Implementation (ITEC Policy 3210)
- Communications Network and Systems Access Security Architecture (ITEC 4210)
- Data Systems Security (PPM 1201.00)
- Development of a Data Administration Program (ITEC Policy 8000)
- DISC Procedures for Off-Site Tape Cartridges (DISC Standard 4453.01)

- Documenting State Data Files Stored on Magnetic Disk and Magnetic Tape (DISC Standard 4412.01)
- Inventory Procedures and Requirements (AC 97-a-003)
- Kansas Geographic Information Systems Metadata Standard (ITEC Policy 5100)
- Project Management (ITEC Policy 2530)
- Security Policy and Procedures for the KANWIN Network (ITEC Policy 4220)
- Standard Two Character Agency/Department Abbreviations and Their Use in Naming Conventions for the DISC Computer Center (DISC Standard 4423.05)
- Technical Architecture Change Management (ITEC Policy 4020)
- United States Government Electronic Commerce Policy - U.S. Department of Commerce
- Year 2000 Asset Readiness Reporting (ITEC Policy 2410)
- Year 2000 Date Data Interchange (ITEC Policy 2412)

20.8.4 Guidelines and Reports

- An Approach to Managing Internet and Intranet Information for Long Term Access and Accountability - Canada
- Digital Signature Guidelines - American Bar Association
- Internet Security Policy: A Technical Guide - National Institute for Standards and Technology (NIST)
- Certification Authority Rating and Trust (CARAT) - National Automated Clearinghouse Associtation (NACHA)
- Digital Signatures & Public Key Infrasturcture (PKI) Guidelines - State of Texas
- BookManager Administration (DISC Guideline 4624.00)
- Copying Vendor Software (DISC Guideline 3607.01)
- Designing and Implementing Recordkeeping Systems (DIRKS Manual) - New South Wales

- Desktop Management: Guidelines for managing electronic documents and directories - New South Wales
- Digital Imaging Guidelines for State Government Records
- Electronic Records Management Guide - U.S. Department of Energy
- Guidelines for Electronic Records Management on State and Federal Agency Websites
- Information Systems Auditing Standards (DISC Guideline 3613.00)
- Kansas Electronic Recordkeeping Strategy: A White Paper
- Kansas Electronic Records Management Guidelines
- Kansas State Records Management Manual
- Managing Electronic Messages as Records - National Archives of Australia
- Managing Electronic Records: A Shared Reponsibility - National Archives of Australia
- Managing Shared Directories and Files - Canada
- Local General Records Retention Schedule
- Local Government Records Management Manual
- Models for Action: Developing Practical Approaches to Electronic Records Management and Preservation - Center for Technology in Government
- Online Reference Manuals (DISC Guideline 4621.02)
- Storage Management Direction (DISC Standard 4463.01)
- Telecommunications Security: Electronic Signature Standardization Report - European Telecommunications Standards Institute
- VMTAPE Expiration Date Coding (DISC Guideline 4625.00)
- WebTrust Principles and Criteria for Certification Authorities - American Institute of Certified Public Accountants and Canadian Institute of Chartered Accountants

20.8.5 Policies of other Countries

- Archiving Web Sites: A policy for keeping web-based records in the Commonwealth Government - Australia
- New South Wales Recordkeeping Metadata Standard (NRKMS) - (proposed)
- Policy for Electronic Recordkeeping in the Commonwealth Government - Australia
- Policy on Electronic Recordkeeping - New South Wales
- Recordkeeping Metadata Standard for Commonwealth Agencies - Australia
- Standard on Full and Accurate Records - New South Wales

20.9 Technical Product & Configuration Information

Category	Emerging	Current	Twilight
Markup tools		• Metabot 2.0 - Watchfire • TagGen - Hiawatha Island Software (HISC)	
Records Management Software	• Accutrac • ForeMost - Provenance Systems • GAIN - Triadd Software • Hummingbird • MASTER TRACK - American Filing Solutions • Open-Text - iRIMS - Web-based records management • OPUS - Thoroughbred Technologies - records processing and management • RetentionManager - Skupsky • RIMS - PSSoftware - owned by Open Text • STAR/RIMS - Cuadra • TRIM Captura - TOWER Software • Versatile Enterprise - Zasio		

20.10 Futures

- Citizen demands for both public access and privacy protection will continue to increase.
- As the state moves toward electronic government, inter-entity and public-private collaboration will increase. Such situations will call for new electronic recordkeeping strategies, since control of information and services will no longer rest solely within individual state entities.
- Policy documents will be increasingly posted online, requiring the state to implement recordkeeping measures to address the resulting accountability exposure.
- More state business will be conducted online, making recordkeeping a vital component of most Internet development projects.
- As an increasingly diverse user base accesses state information and services online, accessibility concerns will become vital. This must include not only the information and services themselves, but public access to open records associated with electronic transactions.
- Electronic transactions will increasinly make use of authentication approaches such as public key-infrastructure (PKI), resulting in the need for long-term management of components such as keys, certificates, algorithms and supporting software. In some cases, the preservation strategy may involve saving preservation copies unencrypted and in others it may involve periodic reauthenticaion of digital objects by trusted custodians.
- Some existing state services will be provided by private entities, requiring the state to address recordkeeping in contractual agreements with those entities.
- On many cases, private entities will make commercial use of pubicly accessible electronic records without seeking permission of the state. The state must develop strategies for addressing these situations.
- The use of software-based electronic records management - both through separate records management applications and integration with existing applications - will increase.
- State information technology projects will address electronic recordkeeping requirements from the beginning of the system lifecycle, up to and including transfer to new systems, when necessary.
- Results of existing research will include more systematic approaches to large-scale data and records preservation. This could potentially make use of some combination of conversion, migration and emulation.

20.11 Organization and Personnel Impact

The Public Records Act (K.S.A. 75-3501 through 75-3518) and the Government Records Preservation Act (K.S.A. 45-401 through 45-413) define the responsibilities of state and local government entities to organize, protect, provide access to, and properly dispose of their records, including the transfer of noncurrent records with enduring value to the Kansas State Historical Society. Cooperation between the entities and the KSHS is even more important with electronic records, because they are more susceptible to loss, inadvertent destruction, mismanagement, and obsolescence. Within entities, cooperation between management, staff who create and handle electronic records, specialists in information system design, and state entity records officers is also essential for the management of electronic records. For these reasons, the KSHS considers the management of electronic records a shared responsibility demanding new partnerships with state entities.

20.11.1 The State Entity's Role

The ability to maintain electronic records and ensure their accessibility over time is highly contingent on how records are created, organized, and maintained in the state entities that create or manage them. Individual state entities are most likely to understand their electronic systems and the specific applications required to maintain the records they contain. As technology changes over time, state entities are also best placed to ensure that records of enduring value are successfully transferred or migrated as systems evolve. In contrast, the KSHS is positioned to provide advice on electronic recordkeeping but without further resource commitment by the state, does not currently have the capacity to manage and maintain a wide range of electronic systems and records applications nor to manage the migration of records to other media and standards over time. Maintenance of most electronic records of long-term value will depend on cooperation between state entities and the State Archives. In order to ensure that records are properly managed, state entities must also cooperate with any other public or private entities with whom they share data for the provision of services.

20.11.1.1 Creation and Maintenance of Electronic Records

Creation and maintenance of reliable and accurate electronic records is the responsibility of program managers, users of computer systems, records officers, and information technology staff who provide technical support and training. End users need to be informed of the policies governing recordkeeping and trained in the use of tools and systems that support electronic records management.

20.11.1.2 Implementation of Records Management Policies

The state entity records officer has responsibility for overseeing the disposition of records, for protecting records with enduring value, and for ensuring that records are not destroyed without authorization of the State Records Board. In extending these responsibilities to include electronic records, it will be necessary for the records officer to participate in studies and analysis of business processes and systems and to participate in the design, monitoring and refining of records storage and retrieval systems. The records officer will also have primary responsibility for applying existing records retention and disposition schedules to electronic records and for submitting new schedules for electronic records that do not have an approved schedule to the State Records Board.

20.11.2 The State Archives' Role

The Kansas State Historical Society is the official State Archives with responsibility to assist state and local entities in the preservation of government records with enduring value (K.S.A. 45-405). While in the past, preservation of such records has been achieved through their physical transfer to the State Archives, preservation of electronic records will depend on closer cooperation with entities. The State Archives will help state entities to identify appropriate maintenance procedures and determine the length of time different types of electronic records should be kept in order to ensure that state entities are not using your resources to maintain records that are no longer needed.

The Archives can help state and local entities to:

- identify the electronic records in state entity custody that are of enduring value,
- identify and obtain authorization to dispose of the electronic records in state entity custody that are not of enduring value,
- identify the metadata that needs to be captured and maintained with electronic records of enduring value if they are to remain identifiable and accessible over time,
- determine the length of time electronic records should be maintained and made accessible in order to meet administrative or archival requirements, and
- develop a means for ensuring the public's right to access to archival electronic records so that you can meet the access provisions of the Kansas Open Records Act (K.S.A. 45-213 through 45-223) while protecting the confidentiality of records exempted from the Open Records Act and other legislation restricting access to records.

In the following limited cases, the State Archives will accept physical custody of electronic records of enduring value:

- It has been demonstrated that the administrative function which the records document has been discontinued by the Kansas legislature and
- there is no successor state entity for the function or activity;
- there is no other state entity or institution which could take custody of the records; or
- the records have a security classification which would preclude them from being stored by another entity or institution.

The Archives enters into an agreement with a state entity to take custody of the electronic records because the alternative arrangements would result in loss of valuable records or represent an uneconomical solution to long-term preservation.

We recognize, however, that there may be equivalent situations where identified temporary value electronic records have to be taken into custody. Each proposed transfer of enduring or temporary value records will be considered on a case-by-case basis (see Section 9.2). Given the wide range of potential formats, volumes, standards of preservation which could be involved and that technology as well as formats and media are subject to constant and rapid change, it is not possible to draw up prescriptive rules governing standards for all proposed transfers of electronic records.

When the Archives has agreed to accept custody of electronic records from a state entity, the Archives and the entity must work together to ensure that the records are transferred to the Archives in an acceptable format and accompanied by the metadata necessary for maintaining access to the records (see Sections 9.3, 9.4 and 9.5).

20.11.3 Shared Responsibilities

Because electronic records management is a shared responsibility, several different organizations in Kansas state government bear some responsibility for implementation, oversight and monitoring. For some activities, this is a logical extension of responsibilities for traditional records management. In some cases, however, this will require new partnerships between administrators, program managers, records officers, and information technology staff in state entities as well as between the entities and the State Archives.

20.11.3.1 System Procurement and Design Standards

The **Information Technology Executive Council (ITEC)** is responsible for approval of information technology policies, project management procedures, the **Kansas Statewide Technical Architecture (KSTA)**, and the strategic information management plan. The KSTA describes the information systems infrastructure that supports the applications used by the State and guides the development of the information systems infrastructure. The **Technical Architecture Review Board** is responsible for keeping the architecture up to date. It evaluates state entity requests for waivers, analyzes projects referred to it to determine architectural compliance, and makes recommendations to ITEC on changes and extensions to the KSTA. To achieve effective integration of recordkeeping requirements and capabilities into new systems, an ongoing partnership between these entities and the State Archives is essential.

The Department of Administration, **Division of Purchases** assists state entities in acquiring goods and services. In order for them to make purchasing decisions that support state electronic recordkeeping needs and for these guidelines to reflect practices that are reasonable from a procurement perspective, the Division of Purchases and State Archives will need to coordinate their activities.

20.11.3.2 Monitoring Records Preservation and Disposal

The State Records Board has primary responsibility for ensuring that important state records are preserved and that other records are disposed of when no longer needed. This includes records in electronic form. Staff of the Records Management Section of the Library and Archives Division, Kansas State Historical Society work with state entity personnel to prepare a retention and disposition schedule, secure state entity approval, and present the schedule to the State Records Board for approval. It is the goal of the Records Management Section to develop retention and disposition schedules for electronic records in conjunction with the design and implementation of new systems whenever possible. ITEC could require that entities have an approved retention and disposition schedule for electronic records before proceeding with the implementation of new systems. Once a schedule has been approved, the state entity records officer is responsible for monitoring its implementation and for recommending revisions to the schedule if requirements or technologies change.